

LOCUS CK230595 1157 bp mRNA linear EST 09-DEC-2003
DEFINITION ILLUMIGEN MCQ 942 Katze MMPL2 Macaca mulatta cDNA 5' similar to human SERPINE2 (Hs.21858), mRNA sequence.
ACCESSION CK230595
VERSION CK230595.1 GI:39636768
KEYWORDS EST.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 1157)
AUTHORS Katze M.G., Thomas M., Korth M., Iadonato S.P. and Magness, C.L.
TITLE Large-scale Rhesus Macaque cDNA Sequencing
JOURNAL Unpublished (2003)
COMMENT Contact: C. Magness
Illumigen Biosciences Inc.
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400
Fax: 2063780408
Email: cmagness@illumigen.com
Sequenced on 2003.10.16. 472 Q20 bases. Assemblies in contig w/ 66 member(s). Contig contains 63 (2.7%) lib members.
PCR PRIMERs
FORWARD: CCTCATTAAGGGAACAAAA
BACKWARD: CACTATAGCGGAATGGGTA
Insert Length: 1157 Std Error: 0.00
Plate: CL060007 row: E column: 08
Seq primer: CCTCATTAAGGGAACAAA
POLYA=Yes.
Location/Qualifiers
1..1157
/organism="Macaca mulatta"
/mol_type="mRNA"
/strain="Indian"
/db_xref="taxon:9544"
/sex="male"
/dev_stage="newborn infant"
/lab_host="E. coli SOLR"
/clone_lib="Katze MMPL2"
/note="Organ: placenta; Vector: Uni-ZAP XR; Site 1: EcoR
I; Site 2: Xho I; Created from Stratagene ZAP-cDNA
Synthesis kit (catalog #200400) and ZAP-CDNA Gigapack III
Gold Cloning Kit (Catalog #200450)"
ORIGIN
Query Match 76.0%; Score 387; DB 14; Length 1157;
Best Local Similarity 94.1%; Pred. No. 7.7e-84;
Matches 413; Conservative 0; Mismatches 25; Indels 1; Gaps 1;
QY 71 TGAATATGATTGGCGAGGAGATCAATATAGGCTAGCCGAGGAGAGTGATACCA 130
DB 1 TGAATATGATTGGCGAGGAGATCAATATAGGCTAGCCGAGGAGAGTGATACCA 60
QY 131 CTTCTGAGCTGATTTGGGCTATGCTGGAGCCGCTGATGAGGAGCTCAGCAAGAGAA 190
DB 61 CTTCTGAGCTGATTTGGGCTATGAGGAGCCGCTGATGAGGAGCTCAGCAAGAGAA 120
QY 191 CCACCAACTGAAGTGGGATCTTGCACTGCTGAGAGAGAGAGATCAGGCTGCA 250
DB 121 CCACCTAACTGAAAGTCAGGATTTACACCTGCTCAGGAGAGAGAGATCAGGCTGCA 180
QY 251 GCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTGCTCAGTCAAGACT 310
DB 181 GCTGAGATTCAAGTGTGATCTGAGCTGAGAGCTGATCTCCAGGAGCAATCTCAGTCAAGACT 240
QY 311 GGGGGTGAATGTGGAATGGTCTGATGACACAGGGGAGATTTCTGCCAAATTCAGAA 370
DB 241 GGGGATGAATCGGAGATGGTCTGATGACACAGGGGAGATTTCTGCCAAATTCAGAA 300
QY 371 TTTAAATGCCAAGAGGTGACAGGCAACACAGGTTTAAATGAAGACAGCTGAAC 430
DB 301 TTTAAATGCCAAGAGGTGACAGGCAACACAGGTTTAAATGAAGACAGCTGAAC 360

QY 431 AACACAAACTCTTTTATCTATAGATATTGCTTAAAAATATCAAAATAACTTTTGA 490
DB 361 AACAC-AAACTCTTTTATCTATAGATATTGCTTAAAAATATGAAATAACTTTTGG 419
QY 491 GCTTTCTCCAAAAA 509
DB 420 GCTTTCTCCAAAAA 438
D59091 394 bp mRNA linear EST 30-AUG-1995
HUM521C10B Clontech human placenta polyA+ mRNA (#6518) Homo sapiens
CDNA clone GEN-521C10 5', mRNA sequence.
D59091
D59091.1 GI:968725
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 394)
AUTHORS Fujiwara, T., Hirano, H., Katagiri, T., Kawai, A., Kuga, Y., Nagata, M., Okuno, S., Ozaki, K., Shimizu, F., Shimada, Y., Shinomiya, H., Takaichi, A., Takeda, S., Watanabe, T., Takahashi, E., Hirai, Y., Maekawa, H., Shin, S. and Nakamura, Y.
TITLE Fujiwara et al. (1995)
JOURNAL Unpublished (1995)
COMMENT Contact: Tsutomu Fujiwara
Otsuka GEN Research Institute
463-10 Kagasuno Kawauchi-cho, Tokushima, Tokushima, 771-01 Japan
Tel: 0886-65-2888
Fax: 0886-37-1035.
Location/Qualifiers
1..394
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GEN-521C10"
/clone_lib="Clontech human placenta polyA+ mRNA (#6518)"
ORIGIN
Query Match 74.7%; Score 380.4; DB 14; Length 394;
Best Local Similarity 99.5%; Pred. No. 2.7e-82;
Matches 392; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 20 GTCGCGTTCTGCTGCTGGACTTTTCTGCTCCACTGAGAGCAGCTGTGTGAATATG 79
DB 1 GTCGCGTTCTGCTGCTGGACTTTTCTGCTCCACTGAGAGCAGCTGTGTGAATATG 60
QY 80 ATTGGCGAGGAGATCAATATAGGCTAGCCGAGGAGAGTGTACCACTCTCTGAG 139
DB 61 ATTGGCGAGGAGATCAATATAGGCTAGCCGAGGAGAGTGTGTACCACTCTCTGAG 120
QY 140 CTGATTGGGCTATGCTGGAGCCGCTGATGAGGAGCTCAGCAAGAGAGCAACCAACT 199
DB 121 CTGATTGGGCTATGCTGGAGCCGCTGATGAGGAGCTCAGCAAGAGAGCAACCAACT 180
QY 200 GAAAGTCGGATCTGCACTGCTGCTGAGAGAGAGATCAGGTCGAGTGTGAGACT 259
DB 181 GAAAGTCGGATCTGCACTGCTGCTGAGAGAGAGATCAGGTCGAGTGTGAGACT 240
QY 260 CAAGTGTGCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGAGCTGGGGTGA 319
DB 241 CAAGTGTGCTGAGCTGGAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGAGCTGGGGTGA 300
QY 320 TGTGGAATGGTCTGATGACAGGGGAGATTTCTGCCAAATTCAGAACTTAAATG 379
DB 301 TGTGGAATGGTCTGATGACAGGGGAGATTTCTGCCAAATTCAGAACTTAAATG 360
QY 380 CCAGAA-GGAGTGTGACAGGCAACCAAGGTTTAA 412

```

Db      361  CCAGAGGGAGGTGACAGGCAACACAGGTTAA 394
|||||
RESULT 15
BX283272
LOCUS   BX283272 NIH MGC 79 Homo sapiens cDNA clone IMAGE4605658, mRNA sequence.
DEFINITION
ACCESSION
VERSION  BX283272
KEYWORDS  EST.
SOURCE    BX283272.1 GI:28847726
ORGANISM  Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 382)
Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,B., Peters,M.,
Radelof,U., Schneider,D. and Korn,B.
Human Unigeneset - RZPD3
Unpublished (2003)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGE958D111343.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human Unigeneset - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
pCMV-M13u, Primer sequence: CGTTGTAACAGCGGCAGT.
FEATURES
source
Location/Qualifiers
1..382
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE958D111343 ; IMAGE:4605658"
/lab_host="DH10B (TI phage-resistant)"
/clone_lib="NIH_MGC_79"
/notes="Organ: placenta; Vector: pDNR-LIB (Clontech);
Site_1: SfiI (ggccgctggcc); Site_2: SfiI
(ggccattatggcc); 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3'
and 3' adaptor sequence:
5'-ATTCTAGCGCGCGCGGCATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.3
kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

```

Search completed: April 26, 2004, 01:58:56
Job time : 928.538 secs

```

Query Match      72.9% Score 371; DB 13; Length 382;
Best Local Similarity 100.0%; Pred. No. 5.3e-80;
Matches 371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTGTGAGGGTGTGAGGTCGGTTCCTGCTGTCTGGACTTTTCTGTCTCCACTGAGA 80
Db 12 GAGTTGTGAGGGTGTGAGGTCGGTTCCTGCTGTCTGGACTTTTCTGTCTCCACTGAGA 71

Qy 61 CGCAGCTGTGTGAATATGATTTGGCGAGGAAGATCAACATATAGCCTAGGCCGAGGAG 120
Db 72 CGCAGCTGTGTGAATATGATTTGGCGAGGAAGATCAACATATAGCCTAGGCCGAGGAG 131

Qy 121 AAGTGTACCACTCTCTGAGCTGATTTGGCCCTATCTCTGAGCCCGGTGATGAGAGCCTCA 180
Db 132 AAGTGTACCACTCTCTGAGCTGATTTGGCCCTATCTCTGAGCCCGGTGATGAGAGCCTCA 191

```

```

Qy 181 GCAAGAGGAACCAACCACTGAAAGTGGGATCTCTCACTCTGTCAGGAGAGAGAAGA 240
Db 192 GCAAGAGGAACCAACCACTGAAAGTGGGATCTCTCACTCTGTCAGGAGAGAGAAGA 251
Qy 241 TCAGGGTGCAGCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCA 300
Db 252 TCAGGGTGCAGCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCA 311
Qy 301 GTCAGAGACTGGGGGTGAATGTGGAATGCTCTGATGACCGAGGAAGATTCTGCCAAA 360
Db 312 GTCAGAGACTGGGGGTGAATGTGGAATGCTCTGATGACCGAGGAAGATTCTGCCAAA 371
Qy 361 ATCAGAAACAAT 371
Db 372 ATCAGAAACAAT 382

```



```

; CURRENT APPLICATION NUMBER: US/09/614.124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 808
; LENGTH: 781
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-614-124B-808

Query Match      42.1%; Score 214.4; DB 4; Length 781;
Best Local Similarity 80.6%; Pred. No. 1.7e-52;
Matches 287; Conservative 0; Mismatches 66; Indels 3; Gaps 3;

QY 155 CTGGAGCCCGGTGATGAGGAGCTCAGCAAGAGGAACACCACAATCTGAAAGTCGGGATCCT 214
DB 426 CAGGAGCCCAAGTAATGGAGGCCCCAAAAGAAGAACAGCAGCTGAAAGTCGGGATCCT 485
QY 215 GCACCTGGTTCAGGAGAGAGAAGATCAGGGTGAGCTGAGACTCAAGTGCCTGACCTG 274
DB 486 ACACCTGGGCAGCAGACAGAAAGATCAGGATACAGCTGAGATCCAGCTCGGCGACATG 545
QY 275 GAAGCTGATCTCCAGGAGCTGTCTCAGTCAAAGACTGGGGGTGAATGTGGAAT-GGTCC 333
DB 546 GAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGGGTTT 605
QY 334 TGATGACCAAGGGGAAGATTCTGCCAAAAATCAGAAATTTAAATGCCAGAGGAGGTGA 393
DB 606 CGCGCTCAAGGTGAAGATAATACCTTAAGAGGAACACACTGTAAATGCCAGAGCAGGTGA 665
QY 394 CAGGCAACACAGGTTTTAAATGAAGACAGAGCTGAAACACACAAAACCTGTTTTATCTAA 453
DB 666 AGAGCAACACCAAGTTTTAAATGAAGACAGAGCTGAAACACGC-AAAGCTGGTTTTATATTA 724
QY 454 GATATTGGACTTTAAAAATATCAAAATAAACTTTTGCAGCTTCTCCAAAAA 509
DB 725 GATATTGGACTT-AAACTATCTCAATAAAGTTTTGCAGCTTTCACCAAAAAA 779

RESULT 6
US-09-671-325-808
; Sequence 808, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fang, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671.325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 808
; LENGTH: 781
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-671-325-808

Query Match      42.1%; Score 214.4; DB 4; Length 781;
Best Local Similarity 80.6%; Pred. No. 1.7e-52;
Matches 287; Conservative 0; Mismatches 66; Indels 3; Gaps 3;

QY 155 CTGGAGCCCGGTGATGAGGAGCTCAGCAAGAGGAACACCACAATCTGAAAGTCGGGATCCT 214
DB 426 CAGGAGCCCAAGTAATGGAGGCCCCAAAAGAAGAACAGCAGCTGAAAGTCGGGATCCT 485
QY 215 GCACCTGGTTCAGGAGAGAGAAGATCAGGGTGAGCTGAGACTCAAGTGCCTGACCTG 274
DB 486 ACACCTGGGCAGCAGACAGAAAGATCAGGATACAGCTGAGATCCAGCTCGGCGACATG 545
QY 275 GAAGCTGATCTCCAGGAGCTGTCTCAGTCAAAGACTGGGGGTGAATGTGGAAT-GGTCC 333
DB 546 GAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGGGTTT 605
QY 334 TGATGACCAAGGGGAAGATTCTGCCAAAAATCAGAAATTTAAATGCCAGAGGAGGTGA 393
DB 606 CGCGCTCAAGGTGAAGATAATACCTTAAGAGGAACACACTGTAAATGCCAGAGCAGGTGA 665
QY 394 CAGGCAACACAGGTTTTAAATGAAGACAGAGCTGAAACACACAAAACCTGTTTTATCTAA 453
DB 666 AGAGCAACACCAAGTTTTAAATGAAGACAGAGCTGAAACACGC-AAAGCTGGTTTTATATTA 724
QY 454 GATATTGGACTTTAAAAATATCAAAATAAACTTTTGCAGCTTCTCCAAAAA 509
DB 725 GATATTGGACTT-AAACTATCTCAATAAAGTTTTGCAGCTTTCACCAAAAAA 779

RESULT 5
US-09-614-124B-808
; Sequence 808, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fang, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9

```

215 GCACCTGGTCCAGGAGAGAGAGATCAGGCTGAGCTGAGACTCAAGTGCCTGACCTG 274
486 ACACCTGGGACAGACAGAGAGATCAGGATCAGCTGAGATCCAGTGGCGACATG 545
275 GAAGCTGATCTCAGGAGCTGTCTGATCAAGAGCTGGGGGTGAATGTGGAAAT-GGTCC 333
546 GAAGTGATCTGCAAGAGCTGATCATCAGTCAACACCGGGGATAAATCTGGATTTGGGTTC 605
334 TGATGACCAAGGGAAGATTCTGCCAAATTCAGAAATCAAGAACTTAAATGCCAGAGGAGGTGA 393
606 CGCGTCAAGGTGAAGATTAATACCTTAAGAGGAACTGTAAATGCCAGAGAGGTGA 665
394 CAGGCAACACAGGTTTAAATGAAGAGAGCTGAAACAAACACAAACTGTTTTATCTAA 453
666 AGAGCAACACAGGTTTAAATGAAGAGAGCTGAAACAAACGC-AAAGCTGGTTTTATTA 724
454 GATATTGACTTAAAAATATCAAAATATAAATTTTGCAGCTTTCTCCAAAAA 509
725 GATATTGACTT-AACTATCTCAATAAAGTTTTCAGCTTTTCCAAAAA 779

RESULT 7
US-09-589-184-808
; Sequence 808, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C8
; CURRENT APPLICATION NUMBER: US/09/589,184
; CURRENT FILING DATE: 2000-06-05
; NUMBER OF SEQ ID NOS: 827
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 808
; LENGTH: 781
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-589-184-808

Query Match 42.1%; Score 214.4; DB 4; Length 781;
Best Local Similarity 80.8%; Pred. No. 1.7e-52;
Matches 287; Conservative 0; Mismatches 66; Indels 3; Gaps 3;

215 CTGAGCCCGGTGATGAGAGCTCTCAGCAAGAGAGAAACCACTGAAAGTCCGGATCCT 214
426 CAGGAGCCAGTAAATGGAGAGCCCAAAAGAGAGAACCACTGAAAGTCCGGATCCT 485
215 GCACCTGGTCCAGGAGAGAGAGATCAGGCTGAGCTGAGACTCAAGTGCCTGACCTG 274
486 ACACCTGGGACAGACAGAGAGATCAGGATCAGCTGAGATCCAGTGGCGACATG 545
275 GAAGCTGATCTCAGGAGCTGTCTGATCAAGAGCTGGGGGTGAATGTGGAAAT-GGTCC 333
546 GAAGTGATCTGCAAGAGCTGATCATCAGTCAACACCGGGGATAAATCTGGATTTGGGTTC 605
334 TGATGACCAAGGGAAGATTCTGCCAAATTCAGAAATCAAGAACTTAAATGCCAGAGGAGGTGA 393
606 CGCGTCAAGGTGAAGATTAATACCTTAAGAGGAACTGTAAATGCCAGAGAGGTGA 665
394 CAGGCAACACAGGTTTAAATGAAGAGAGCTGAAACAAACACAAACTGTTTTATCTAA 453
666 AGAGCAACACAGGTTTAAATGAAGAGAGCTGAAACAAACGC-AAAGCTGGTTTTATTA 724
454 GATATTGACTTAAAAATATCAAAATATAAATTTTGCAGCTTTCTCCAAAAA 509

725 GATATTGACTT-AACTATCTCAATAAAGTTTTCAGCTTTTCCAAAAA 779

RESULT 8
US-09-702-705-69
; Sequence 69, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 69
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-69

Query Match 42.0%; Score 214; DB 4; Length 399;
Best Local Similarity 80.3%; Pred. No. 1.6e-52;
Matches 286; Conservative 1; Mismatches 66; Indels 3; Gaps 3;

155 CTGAGCCCGGTGATGAGAGCTCTCAGCAAGAGAGAACCACTGAAAGTCCGGATCCT 214
44 CAGGAGCCAGTAAATGGAGAGCCCAAAAGAGAGAACCACTGAAAGTCCGGATCCT 103
215 GCACCTGGTCCAGGAGAGAGAGATCAGGCTGAGCTGAGACTCAAGTGCCTGACCTG 274
104 ACACCTGGGACAGACAGAGAGAGATCAGATACAGCTGAGATCCAGTGGCGACATG 163
275 GAAGCTGATCTCAGGAGCTGTCTGATCAAGAGCTGGGGGTGAATGTGGAAAT-GGTCC 333
164 GAAGTGATCTGCAAGAGCTGATCATCAGTCAACACCGGGGATAAATCTGGATTTGGGTTC 223
334 TGATGACCAAGGGAAGATTCTGCCAAATTCAGAACTTAAATGCCAGAGAGGTGA 393
224 CGCGTCAAGGTGAAGATTAATACCTTAAGAGGAACTGTAAATGCCAGAGAGGTGA 283
394 CAGGCAACACAGGTTTAAATGAAGAGAGCTGAAACAAACACAAACTGTTTTATCTAA 453
284 AGAGCAACCAAGTTTAAATGAAGAGAGCTGAAACAAACGC-AAAGCTGGTTTTATTA 342
454 GATATTGACTTAAAAATATCAAAATATAAATTTTGCAGCTTTCTCCAAAAA 509
343 GATATTGACTT-AACTATCTCAATAAAGTTTTCAGCTTTTCCAAAAA 397

RESULT 9
US-09-736-457-69
; Sequence 69, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane

APPLICANT: Fan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 69
LENGTH: 399
TYPE: DNA
ORGANISM: Homo sapien
US-09-736-457-69

Query Match 42.0%; Score 214; DB 4; Length 399;
Best Local Similarity 80.3%; Pred. No. 1.6e-52;
Matches 286; Conservative 1; Mismatches 66; Indels 3; Gaps 3;
QY 155 CTGGAGCCCGGTGATGAGGAGCCTCAGCAAGAGAGGAAACACCAACTGAAAGTCGGGATCCT 214
DB 44 CAGGAGCCCAAGTAATGGAGAGGCCCAAAAGAGAGAACACAGAGCTGAAAGTCGGGATCCT 103
QY 215 GCACCTGTGTCAGGAGAGAGAAATCAGGTCAGGTCGAGCTGAGACTCAAGTCGCTGACCTG 274
DB 104 ACACCTGGGCGAGCAGACAGAGAGATCAGAGTACAGCTGAGATCCCACTGGCGGACATG 163
QY 275 GAAGCTGATCTCCAGGAGCTGTCTCAGTCAAAAGACTGGGGGTGAATGTGGAAT -GGTCC 333
DB 164 GAAGGTGATCTGCAGAGCTGCAATCAGTCAACACCGGGGATAATCTGGATTTGGGTTTC 223
QY 334 TGATGACCGAGGGAAGTCTTGCACAAATCAGAACAAATTTAAATCCAGAGAGAGGTGA 393
DB 224 CGCGGTCAAGGTGAGAGTAATACCTAAAGAGGAACTCTGTAATAATCCAGAGAGGTGA 283
QY 394 CAGGCAACACACAGGTTTAAATGAAGCAAGCTGAAACACACAAACTGTTTATCTAA 453
DB 284 AGAGCAACACACAGTTTAAATGAAGCAAGCTGAAACACACGC -AAGCTGTTTATTA 342
QY 454 GATATTGACTTAAATAATATCAAAATATACTTTTGAGCTTTTCCAAAAA 509
DB 343 GATATTGACTT-AAACTATCTCAATAAAGTTTTCAGCTTTTCAACCAAAAAA 397

RESULT 10
US-09-614-124B-69
Sequence 69, Application US/09614124B
Patent No. 6630574
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C9
CURRENT APPLICATION NUMBER: US/09/614,124B
CURRENT FILING DATE: 2001-07-11
NUMBER OF SEQ ID NOS: 1668
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 69
LENGTH: 399
TYPE: DNA
ORGANISM: Homo sapien
US-09-614-124B-69

Query Match 42.0%; Score 214; DB 4; Length 399;
Best Local Similarity 80.3%; Pred. No. 1.6e-52;
Matches 286; Conservative 1; Mismatches 66; Indels 3; Gaps 3;

QY 155 CTGGAGCCCGGTGATGAGGAGCCTCAGCAAGAGAGGAAACACCAACTGAAAGTCGGGATCCT 214
DB 44 CAGGAGCCCAAGTAATGGAGAGGCCCAAAAGAGAGAACACAGAGCTGAAAGTCGGGATCCT 103
QY 215 GCACCTGTGTCAGGAGAGAGAAATCAGGTCAGGTCGAGCTGAGACTCAAGTCGCTGACCTG 274
DB 104 ACACCTGGGCGAGCAGACAGAGAGATCAGAGTACAGCTGAGATCCCACTGGCGGACATG 163
QY 275 GAAGCTGATCTCCAGGAGCTGTCTCAGTCAAAAGACTGGGGGTGAATGTGGAAT -GGTCC 333
DB 164 GAAGGTGATCTGCAGAGCTGCAATCAGTCAACACCGGGGATAATCTGGATTTGGGTTTC 223
QY 334 TGATGACCGAGGGAAGTCTTGCACAAATCAGAACAAATTTAAATCCAGAGAGAGGTGA 393
DB 224 CGCGGTCAAGGTGAGAGTAATACCTAAAGAGGAACTCTGTAATAATCCAGAGAGGTGA 283
QY 394 CAGGCAACACACAGGTTTAAATGAAGCAAGCTGAAACACACAAACTGTTTATCTAA 453
DB 284 AGAGCAACACACAGTTTAAATGAAGCAAGCTGAAACACACGC -AAGCTGTTTATTA 342
QY 454 GATATTGACTTAAATAATATCAAAATATACTTTTGAGCTTTTCCAAAAA 509
DB 343 GATATTGACTT-AAACTATCTCAATAAAGTTTTCAGCTTTTCAACCAAAAAA 397

RESULT 11
US-09-671-325-69
Sequence 69, Application US/09671325
Patent No. 6667154
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C12
CURRENT APPLICATION NUMBER: US/09/671,325
CURRENT FILING DATE: 2000-09-26
NUMBER OF SEQ ID NOS: 1825
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 69
LENGTH: 399
TYPE: DNA
ORGANISM: Homo sapien
US-09-671-325-69

Query Match 42.0%; Score 214; DB 4; Length 399;
Best Local Similarity 80.3%; Pred. No. 1.6e-52;
Matches 286; Conservative 1; Mismatches 66; Indels 3; Gaps 3;
QY 155 CTGGAGCCCGGTGATGAGGAGCCTCAGCAAGAGAGGAAACCAACTGAAAGTCGGGATCCT 214
DB 44 CAGGAGCCCAAGTAATGGAGAGGCCCAAAAGAGAGAACACAGAGCTGAAAGTCGGGATCCT 103
QY 215 GCACCTGTGTCAGGAGAGAGAGATCAGGTCGAGCTGAGACTCAAGTCGCTGACCTG 274
DB 104 ACACCTGGGCGAGCAGACAGAGAGATCAGAGTACAGCTGAGATCCCACTGGCGGACATG 163
QY 275 GAAGCTGATCTCCAGGAGCTGTCTCAGTCAAAAGACTGGGGGTGAATGTGGAAT -GGTCC 333
DB 164 GAAGGTGATCTGCAGAGCTGCAATCAGTCAACACCGGGGATAATCTGGATTTGGGTTTC 223
QY 334 TGATGACCGAGGGAAGTCTTGCACAAATCAGAACAAATTTAAATCCAGAGAGAGGTGA 393
DB 224 CGCGGTCAAGGTGAGAGTAATACCTAAAGAGGAACTCTGTAATAATCCAGAGAGGTGA 283

[illegible]

RESULT 12

```

US-09-589-184-69
; Sequence 69, Application US/09589184
; Patent NO. 6686447
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C8
; CURRENT APPLICATION NUMBER: US/09/589,184
; CURRENT FILING DATE: 2000-06-05
; NUMBER OF SEQ ID NOS: 827
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 69
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-589-184-69

```

RESULT 13

```

RESULT 13
US-09-163-748C-1
; Sequence 1, Application US/09163748C
; Patent No. 6509172
; GENERAL INFORMATION:
; APPLICANT: Debacker, Oliver
; APPLICANT: Van den Eynde, Benoit
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Which Encode Member Of The Gage

```

```

; TITLE OF INVENTION: Family, The Proteins Encoded, And Uses Thereof
;
; FILE REFERENCE: LUD 5558
;
; CURRENT APPLICATION NUMBER: US/09/163,748C
;
; CURRENT FILING DATE: 1998-09-30
;
; NUMBER OF SEQ ID NOS: 27
;
; SEQ ID NO 1
;
; LENGTH: 528
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
; FEATURE:
;
; US-09-163-748C-1

```

RESULT 14

US-08-370-648-14
Sequence 14, Application US/08370648
Patent No. 5648226
GENERAL INFORMATION:
APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
APPLICANT: Boon-Falleur, Thierry;
TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen, and Uses Thereof
TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/370,648
FILING DATE: 10-JANUARY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250,162
FILING DATE: 27-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/096,039
FILING DATE: 22-JULY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5648226man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5323.2
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 538 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-370-648-14

Query Match 41.1%; Score 209; DB 1; Length 538;
Best Local Similarity 69.2%; Pred. No. 5.3e-51;
Matches 352; Conservative 0; Mismatches 135; Indels 22; Gaps 4;
QY 20 GTCCGCTTCCTGCTGCTGAGCTTTTCTGCTCCACTGAGAGCGAGCTGTGTGAATATG 79
DB 27 GTGTGTTCTCTGCTGCTGAGCTTTTCTCTACTGAGATTCATCTGTGTGAATATG 86
QY 80 ATTTGCGGAGAGATCAACATATAGCCCTAGCCGAGAGAGTGTACCACTCTCTGAG 139
DB 87 AGTTGCGGAGAGATCAACATATAGCCCTAGCCGAGAGTGTACCACTCTCTGAG 146
QY 140 CTGATTTGGGCTTATGC-----TGGAGCCCGGTGATGAGGAG-----CCTCAG 181
DB 147 ATGATTTGGGCTTATGCGGCCGAGAGCTTCTGATGATGAGTGAACACCACTGAA 206
QY 302 TCAAGAGCTGGGGTGAATGTGAATGTTCTGATGACCGGGAAGATTTCTCCAAA 361
DB 324 CCACAGACTGGGTGTGAGTGTGAATGTTCTGATGGGCGAGAGATGGAACCGGCAAT 383
QY 362 TCAGAACATTTAAATATGCGAGAGAGTGTGACAGGCAACCACTGTTTAAATGAGACA 421
DB 384 CCAGAGGAGGTGAACAGCTGAGAGAGTGTGAAGAGTGAAGAGCAATCACAGTGTAAAGAGACA 443
QY 422 AGCTGAACACACAAAAGTGT-TTTTATCTAAGATATTTGACTTAAATATCAAAATA 480
DB 444 CGTTGAATATGATGCGAGGCTGCTCTATGTTGAAATTTGTTCAATTAATTTCTCCAAATA 503
QY 481 AACTTTTGAGCTTTCTTCCAAAAA 509
DB 504 AGCATTACAGCTTCTGCAAGAAAAA 532

RESULT 15
US-08-531-662B-14
; Sequence 14, Application US/08531662B
; Patent No. 5858689

GENERAL INFORMATION:

APPLICANT: Van der Bruggen, Pierre;
APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
APPLICANT: Boon-Palleur, Thierry
TITLE OF INVENTION: Isolated, Truncated Nucleic
TITLE OF INVENTION: Acid Molecules Which Code For GAGE Tumor
TITLE OF INVENTION: Rejection Antigen, The Tumor Rejection Antigen,
TITLE OF INVENTION: and Uses Thereof
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/531,662B
FILING DATE: 21-September-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/370,648
FILING DATE: 10-January-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250,162
FILING DATE: 27-May-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/096,039
FILING DATE: 22-July-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5858689man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5323.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 538 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-531-662B-14

Query Match 41.1%; Score 209; DB 2; Length 538;
Best Local Similarity 69.2%; Pred. No. 5.3e-51;
Matches 352; Conservative 0; Mismatches 135; Indels 22; Gaps 4;
QY 20 GTCCGCTTCCTGCTGCTGAGCTTTTCTGCTCCACTGAGAGCGAGCTGTGTGAATATG 79
DB 27 GTGTGTTCTCTGCTGCTGAGCTTTTCTCTACTGAGATTCATCTGTGTGAATATG 86
QY 80 ATTTGCGGAGAGATCAACATATAGCCCTAGCCGAGAGAGTGTACCACTCTCTGAG 139
DB 87 AGTTGCGGAGAGATCAACATATAGCCCTAGCCGAGAGTGTACCACTCTCTGAG 146
QY 140 CTGATTTGGGCTTATGC-----TGGAGCCCGGTGATGAGGAG-----CCTCAG 181
DB 147 ATGATTTGGGCTTATGCGGCCGAGAGCTTCTGATGATGAGTGAACACCACTGAA 206
QY 182 CAAGAGAACCAACAACTGAAAGTCTGTGCACTCTGCTGAGAGAGAGAT 241
DB 207 GAAGGGGAACCAAGCACTCAACGCTCAGGATCTCTGAGCTGCTCAGGAG---GGAGAGGAT 263
QY 242 CAGGTCAGCTGAGACTCAAGTGCCTGAGAGAGTGTATCTCCAGAGCTGTCTCAG 301
DB 264 GAGGAGCATCTGAGGTCAAGGCGGAGAGCTTGAAGCTCATAGCCAGGACAGGCTCAC 323

Qy	302	TCAAAGACTGGGGTGAATGTGGAATGGTCTCTGATGACCCAGGGGAAGATTCTGCCAAAA	361
Db	324	CCACAGACTGGGTGTGAGTGTGAGATGGTCTCTGATGGGCGAGAGATGGACCGCCCAAT	383
Qy	362	TCAGAACAAATTTAAATGCCAGAGGAGGTGACAGGCAACCCACAGGTTTAAATGAAGACA	421
Db	384	CCAGAGGAGGTGAAGACGCCCTGAAGAAGGTGAAGCAATCACAGTGTAAAGAGAGACA	443
Qy	422	AGCTGAAACAAACACAAACTGT-TTTTATCTAAGATATTGACTTAAATAATATCAAAATA	480
Db	444	CGTTGAATGATGCAGGCTGCTCTATGTTGGAAATTGTTTCATTAAATTTCTCCCAATA	503
Qy	481	AACTTTGCAGCTTCTCCAAAAA	509
Db	504	AAGCTTTACAGCCTTCTGCAAGAAAAA	532

Search completed: April 26, 2004, 02:10:45
Job time : 26.2508 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2004, 23:47:42 ; Search time 135,244 Seconds
(without alignments)
15988.422 Million cell updates/sec

Title: US-10-051-835-11
Perfect score: 509
Sequence: 1 gagggtgaggtgtgaggg.....agcttttcccaaaaaaaa 509

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 337863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_24Jan04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002s.*
7: Geneseqn2003as.*
8: Geneseqn2003bs.*
9: Geneseqn2003cs.*
10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	509	100.0	509	7 ABX77605	Abx77605 Different
2	509	100.0	509	8 ACD42232	Ac42232 Human GAG
3	509	100.0	509	9 ADC24846	Adc24846 Human GAG
4	504.2	99.1	520	7 ABT15737	Abt15737 Human can
5	483.4	95.0	659	4 AAI58744	Aai58744 Human pol
6	483.4	95.0	659	8 ADB48724	Adb48724 Novel hum
7	480.6	94.4	532	4 AAI60530	Aai60530 Human pol
8	421.4	82.8	503	5 AAS69484	Aas69484 DNA encod
9	393.8	77.4	661	5 AAS69486	Aas69486 DNA encod
10	337.4	66.3	642	6 ABA93876	Ab93876 Human G p
11	337.4	66.3	673	4 AAK51918	Aak51918 Human pol
12	337.4	66.3	756	7 ABT15738	Abt15738 Human can
13	322.8	63.4	611	4 AAD14983	Aad14983 Human NOV
14	292.6	57.5	580	4 AAF59637	Aaf59637 Human cel
15	290.2	57.0	669	8 ACD42228	Ac42228 Human GAG
16	237.4	46.6	334	7 ABT15743	Abt15743 Human can
17	233	45.8	665	4 AAK52902	Aak52902 Human pol
18	214.4	42.1	626	7 ACC51056	Acc51056 Human bla
19	214.4	42.1	626	7 ABX76178	Abx76178 Lung can
20	214.4	42.1	626	7 ACC49534	Acc49534 Tumour-as
21	214.4	42.1	626	9 ADB75626	Adb75626 Prostate
22	214.4	42.1	637	6 ABX50912	Abx50912 XAGE-1 ge
23	214.4	42.1	762	5 AAB64751	Aab64751 Human sec

24	214.4	42.1	781	5 AAF68861	Aaf68861 Human lun
25	214.4	42.1	781	6 ABK38772	Abk38772 cDNA enco
26	214.4	42.1	781	7 ACA11101	Ac11101 Human lun
27	214.4	42.1	781	7 ACA02287	Ac02287 Lung can
28	214	42.0	399	5 AAF68151	Aaf68151 Human lun
29	214	42.0	399	6 ABK38062	Abk38062 cDNA enco
30	214	42.0	399	7 ACA10391	Ac10391 Human lun
31	214	42.0	399	7 ABX99342	Abx99342 Lung can
32	214	42.0	399	10 ADE72125	Ad72125 Human lun
33	211.4	41.5	463	6 ABA92217	Ab92217 Melanoma
34	210.6	41.4	528	7 ADA15801	Ad15801 Human GAG
35	209	41.1	538	2 AAX90519	Aax90519 GAGE-2 tu
36	209	41.1	1345	4 AAS60826	Aas60826 Human can
37	207	40.7	530	7 ABZ20463	Abz20463 GAGE-2 fu
38	205	40.3	541	2 AAV18719	Aav18719 cDNA enco
39	202.8	39.8	539	2 AAV18721	Aav18721 cDNA enco
40	202.4	39.8	540	2 AAX90523	Aax90523 GAGE-6 tu
41	202.4	39.8	540	2 AAX90521	Aax90521 GAGE-4 tu
42	201.8	39.6	535	2 AAV18717	Aav18717 cDNA enco
43	201.6	39.6	750	5 AAS91235	Aas91235 DNA encod
44	201.2	39.5	532	2 AAV18720	Aav18720 cDNA enco
45	200.8	39.4	532	2 AAX90522	Aax90522 GAGE-5 tu

ALIGNMENTS

RESULT 1
ABX77605
ID ABX77605 standard; cDNA; 509 BP.
AC ABX77605;
XX
DT 09-APR-2003 (first entry)
XX
DE Differentially expressed breast cancer associated cDNA #100.
XX
KW Breast cancer; differential gene expression; BC-cDNA;
KW breast cancer diagnosis; breast cancer monitoring;
KW breast cancer treatment; breast cancer staging; gene; ss.
XX
OS Unidentified.
XX
PN US2002156263-A1.
XX
PD 24-OCT-2002.
XX
PF 04-OCT-2001; 2001US-00974298.
XX
PR 05-OCT-2000; 2000US-0238331P.
XX
PA (CHEN/) CHEN H.
XX
PI Chen H;
XX
DR WPI; 2003-182653/18.
XX
PT New cDNAs, which are differentially expressed in (metastatic) breast
PT cancer useful for diagnosing or staging, breast cancer, or for monitoring
PT the treatment of breast cancer in an individual.
XX
PS Claim 1; SEQ ID NO 126; 30pp; English.
XX
CC The invention describes a combination of cDNAs (designated BC-cDNAs),
CC which are differentially expressed in breast cancer. The combination
CC includes 152 cDNA sequences, or their complements. The protein encoded by
CC any of these BC-cDNAs is useful for screening several molecules or
CC compounds to identify at least one ligand that specifically binds the
CC protein, producing or preparing polyclonal or monoclonal antibodies, or
CC purifying antibodies from a sample. The antibodies, which specifically
CC bind the protein differentially expressed in breast cancer is useful for
CC detecting the expression of a protein in a sample. The BC-cDNAs are also
CC useful for diagnosing, monitoring the treatment of, or staging, breast

CC cancer. This sequence represents a differentially expressed breast cancer
CC associated cDNA. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?DocID=20020156263
XX
SQ Sequence 509 BP; 158 A; 100 C; 139 G; 112 T; 0 U; 0 Other;

Query Match 100.0%; Score 509; DB 7; Length 509;
Best Local Similarity 100.0%; Pred. No. 1.1e-133;
Matches 509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGTTGTGAGGGTGTGAGGGTCCGCTTCTGCTGTCTGCTGACTTTTCTCTCCACTGAGA 60
Db 1 GAGTTGTGAGGGTGTGAGGGTCCGCTTCTGCTGTCTGCTGACTTTTCTCTCCACTGAGA 60
QY 61 CGCAGCTGTGTGAATATGATTTGGCGAGAGATCAACATATAGCCCTAGSCCGAGGAG 120
Db 61 CGCAGCTGTGTGAATATGATTTGGCGAGAGATCAACATATAGCCCTAGSCCGAGGAG 120
QY 121 AAGTGTACCACTCCCTGAGCTGATTGGGCTATGCTGGAGCCCGGTGATGAGGAGCTCA 180
Db 121 AAGTGTACCACTCCCTGAGCTGATTGGGCTATGCTGGAGCCCGGTGATGAGGAGCTCA 180
QY 181 GCAAGAGAACCAACCACTGAAAGTCGGATCTGCTGAGCCCGGTGATGAGGAGCTCA 240
Db 181 GCAAGAGAACCAACCACTGAAAGTCGGATCTGCTGAGCCCGGTGATGAGGAGCTCA 240
QY 241 TCAGGGTGCAGCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTCTCA 300
Db 241 TCAGGGTGCAGCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTCTCA 300
QY 301 GTCAAGAGACTGGGGGTGAATGTGGAAATGGTCTCTGATGACACGAGGAGAAATCTGCCAAA 360
Db 301 GTCAAGAGACTGGGGGTGAATGTGGAAATGGTCTCTGATGACACGAGGAGAAATCTGCCAAA 360
QY 361 ATCAGAACCAATTTAAATGCCAGAGAGGTGACAGGCAACCAAGGTTTAAATGAGAC 420
Db 361 ATCAGAACCAATTTAAATGCCAGAGAGGTGACAGGCAACCAAGGTTTAAATGAGAC 420
QY 421 AAGCTGAAACCAACACAAAAGTGTCTTATCTAAGATATTTGACTTAAATAATCAAAATA 480
Db 421 AAGCTGAAACCAACACAAAAGTGTCTTATCTAAGATATTTGACTTAAATAATCAAAATA 480
QY 481 AACTTTGAGCTTTCTCCAAAAA 509
Db 481 AACTTTGAGCTTTCTCCAAAAA 509

RESULT 2
ACD42232
ID ACD42232 standard; cDNA; 509 BP.
XX
AC ACD42232;
XX
AC ACD42232;
XX
DT 05-SEP-2003 (first entry)
XX
DE Human GAGE-8 cDNA Incyte 064516CB1.
XX
XX Human; ss; DNA methylation; cancer; colon cancer.
XX
OS Homo sapiens.
XX
PN US2003013099-A1.
XX
XX 16-JAN-2003.
XX
XX 07-MAR-2002; 2002US-00093766.
XX
XX 19-MAR-2001; 2001US-0277380P.
XX
XX (LASEK/) LASEK A K W.
XX
XX (JONES/) JONES D A.

PA (KARP/) KARP A R.
XX
PI Lasek AKW, Jones DA, Karpf AR;
XX
DR WPI; 2003-503249/47.
DR P-PSDB; ABO25121.
XX
PT New combination comprising cDNAs that are expressed in a disorder or
PT process associated with DNA methylation, useful for diagnosing, staging,
PT treating or monitoring treatment of cancer, e.g. colon cancer.
XX
PS Claim 2; Page 43-44; 66pp; English.
XX
The invention relates to a combination comprising cDNAs which are
CC expressed in a disorder or process associated with DNA methylation. The
CC combination and cDNAs are useful for diagnosing, staging, treating or
CC monitoring treatment of cancer, e.g. colon cancer and for detecting
CC changes in expression of genes encoding proteins that are associated with
CC DNA methylation. The protein is useful for screening molecules or
CC compounds to identify at least one ligand that binds to the protein and
CC for producing an antibody. The present sequence represents a cDNA
CC expressed in a disorder or process associated with DNA methylation
XX
SQ Sequence 509 BP; 158 A; 100 C; 139 G; 112 T; 0 U; 0 Other;

Query Match 100.0%; Score 509; DB 8; Length 509;
Best Local Similarity 100.0%; Pred. No. 1.1e-133;
Matches 509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGTTGTGAGGGTGTGAGGGTCCGCTTCTGCTGTCTGGAATTTTCTGCTCCACTGAGA 60
Db 1 GAGTTGTGAGGGTGTGAGGGTCCGCTTCTGCTGTCTGGAATTTTCTGCTCCACTGAGA 60
QY 61 CGCAGCTGTGTGAATATGATTTGGCGAGAGATCAACATATAGCCCTAGSCCGAGGAG 120
Db 61 CGCAGCTGTGTGAATATGATTTGGCGAGAGATCAACATATAGCCCTAGSCCGAGGAG 120
QY 121 AAGTGTACCACTCCCTGAGCTGATTGGGCTATGCTGGAGCCCGGTGATGAGGAGCTCA 180
Db 121 AAGTGTACCACTCCCTGAGCTGATTGGGCTATGCTGGAGCCCGGTGATGAGGAGCTCA 180
QY 181 GCAAGAGAACCAACCACTGAAAGTCGGATCTGCTGAGCCCGGTGATGAGGAGCTCA 240
Db 181 GCAAGAGAACCAACCACTGAAAGTCGGATCTGCTGAGCCCGGTGATGAGGAGCTCA 240
QY 241 TCAGGGTGCAGCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTCTCA 300
Db 241 TCAGGGTGCAGCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTCTCA 300
QY 301 GTCAAGAGACTGGGGGTGAATGTGGAAATGGTCTCTGATGACACGAGGAGAAATCTGCCAAA 360
Db 301 GTCAAGAGACTGGGGGTGAATGTGGAAATGGTCTCTGATGACACGAGGAGAAATCTGCCAAA 360
QY 361 ATCAGAACCAATTTAAATGCCAGAGAGGTGACAGGCAACCAAGGTTTAAATGAGAC 420
Db 361 ATCAGAACCAATTTAAATGCCAGAGAGGTGACAGGCAACCAAGGTTTAAATGAGAC 420
QY 421 AAGCTGAAACCAACACAAAAGTGTCTTATCTAAGATATTTGACTTAAATAATCAAAATA 480
Db 421 AAGCTGAAACCAACACAAAAGTGTCTTATCTAAGATATTTGACTTAAATAATCAAAATA 480
QY 481 AACTTTGAGCTTTCTCCAAAAA 509
Db 481 AACTTTGAGCTTTCTCCAAAAA 509

RESULT 3
ADC24646
ID ADC24646 standard; cDNA; 509 BP.
XX
AC ADC24646;
XX
DT 18-DEC-2003 (first entry)

QY	241	TCAGGGTCGACTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCA	300
DB	241	TCAGGGTCGACTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCA	300
QY	301	GTCAAAGACTCGGGGTGAATGTGGAATGCTCTGTATGACCCAGGGGAAGATTCTGCCAAA	360
DB	301	GTCAAAGACTCGGGGTGAATGTGGAATGCTCTGTATGACCCAGGGGAAGATTCTGCCAAA	360
QY	361	ATCAGAACAAATTTAAATGCCAGNAGAGGTGACAGGCAACACACAGGTTTAAATGAGAC	420
DB	361	ATCAGAACAAATTTAAATGCCAGNAGAGGTGACAGGCAACACACAGGTTTAAATGAGAC	420
QY	421	AAAGCTGAAACCAACACAAAACCTGTTTTTATCTAAGATATTTGACTTAAAAATATCAAATA	480
DB	421	AAAGCTGAAACCAACACAAAACCTGTTTTTATCTAAGATATTTGACTTAAAAATATCAAATA	480
QY	481	AACCTTTTCAGCTTCTCCAAAAA	509
DB	481	AACCTTTTCAGCTTCTCCAAAAA	509
RESULT 4			
ABT15737			
ID	ABT15737 standard; DNA; 520 BP.		
AC	ABT15737;		
XX			
DT	28-MAR-2003 (first entry)		
DE	Human cancer/testis antigen coding sequence - SEQ ID No 38.		
KW	Human; gene; ds; gene therapy; vaccine; cancer; cancer/testis antigen;		
KM	CT antigen.		
XX			
XX	Homo sapiens.		
OS			
XX	WO200278526-A2.		
PN			
PD	10-OCT-2002.		
XX			
XX	29-MAR-2002; 2002WO-US009808.		
PF			
PR	30-MAR-2001; 2001US-0280718P.		
PR	20-APR-2001; 2001US-0285154P.		
PR	05-OCT-2001; 2001US-0327432P.		
PR	22-JAN-2002; 2002US-00054683.		
XX			
PA	{LUDW-} LUDWIG INST CANCER RES.		
PA	{CORR } CORNELL RES FOUND INC.		
XX			
PI	Old LJ, Scanlan MJ, Chen Y;		
XX			
DR	WPI; 2003-040608/03.		
DR	P-PSDB; ABJ19256.		
XX			
PT	Diagnosing cancer comprises contacting a biological sample isolated from		
PT	a subject with an agent that specifically binds to a nucleic acid		
PT	molecule, its expression product or fragment or an antibody that binds to		
PT	the product or fragment.		
XX			
PS	Claim 7; Page 145; 155pp; English.		
XX			
CC	The invention comprises a method for diagnosing cancer, the method		
CC	involves detecting the DNA or protein sequences of human cancer/testis		
CC	(CT) antigens that are disclosed in the invention. The method of the		
CC	invention is useful for detecting/diagnosing, treating and monitoring a		
CC	cancer or condition characterised by the expression of a human CT		
CC	antigen. The present DNA sequence encodes a human CT antigen of the		
CC	invention		
XX			
SQ	Sequence 520 BP; 167 A; 100 C; 140 G; 113 T; 0 U; 0 Other;		

Query Match	99.1%;	Score 504.2;	DB 7;	Length 520;	
Best Local Similarity	99.4%;	Pred. No. 2.4e-132;			
Matches 506;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;	
QY	1	GAGTTGTGAGGTGTGAGGGTGGCGTCTCTGCTGTCTGAGCTTTTCTGTCCCACTGAGA	60		
Db	1	GAGTTGTGAGGTGTGAGGGTGGCGTCTCTGCTGTCTGAGCTTTTCTGTCCCACTGAGA	60		
QY	61	CCGAGCTGTGTGAATATGATTTGGCGAGGAAGATCAACATATAGGCTAGGCCGAGGAG	120		
Db	61	CCGAGCTGTGTGAATATGATTTGGCGAGGAAGATCAACATATAGGCTAGGCCGAGGAG	120		
QY	121	AAGTGTACCACTCTGAGCTGATTTGGGCTATGCTGAGCGCGGTGATGAGGACCTCA	180		
Db	121	AAGTGTACCACTCTGAGCTGATTTGGGCTATGCTGAGCGCGGTGATGAGGACCTCA	180		
QY	181	GCAAGAGAACCAACCACTGAAGTCGGGATCCTGCACCTGCTGAGGAGAGAAGA	240		
Db	181	GCAAGAGAACCAACCACTGAAGTCGGGATCCTGCACCTGCTGAGGAGAGAAGA	240		
QY	241	TCAGGGTGCAGTGTGAGCTCAAGTGCCTGGAAGTGAATGCTGATGACAGGGAAGATTCTGCCAAA	300		
Db	241	TCAGGGTGCAGTGTGAGCTCAAGTGCCTGGAAGTGAATGCTGATGACAGGGAAGATTCTGCCAAA	300		
QY	301	GTCAGAGCTGGGGTGAATGTGGAATGGTCTCTGATGACAGGGAAGATTCTGCCAAA	360		
Db	301	GTCAGAGCTGGGGTGAATGTGGAATGGTCTCTGATGACAGGGAAGATTCTGCCAAA	360		
QY	361	ATCAGAGCAATTTAAATCCAGAGAGGAGTGCACAGGCAACACAGGTTTAAATGAAGAC	420		
Db	361	ATCAGAGCAATTTAAATCCAGAGAGGAGTGCACAGGCAACACAGGTTTAAATGAAGAC	420		
QY	421	AAGCTGAAACAAACACAAAACCTGTTTATCTAAGATATTTGACTTTAAAAATATCAAAATA	480		
Db	421	AAGCTGAAACAAACCAAAACCTGTTTATCTAAGATATTTGACTTTAAAAATATCAAAATA	480		
QY	481	AACTTTTGAGCTTTCTCCAAAAA	509		
Db	481	AACTTTTGAGCTTTCTCCAAAAA	509		
RESULT 5					
AAI58744					
ID	AAI58744	standard; cDNA; 659 BP.			
XX	AAI58744;				
AC	AAI58744;				
XX	22-OCT-2001	(first entry)			
DT	Human polynucleotide	SEQ ID NO 947.			
DE	Human;	nootropic; immunosuppressant; cytostatic; gene therapy; cancer;			
XX	peripheral nervous system; neuropathy; central nervous system; CNS;				
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;				
KW	anyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;				
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;				
KW	leukaemia; ss.				
XX	Homo sapiens.				
OS	WO200153312-A1.				
PN	26-JUL-2001.				
XX	26-DEC-2000;	2000WO-US034263.			
XX	23-DEC-1999;	99US-00471275.			
PR	21-JAN-2000;	2000US-00488725.			
PR	25-APR-2000;	2000US-00552317.			
PR	20-JUN-2000;	2000US-00598042.			
PR	19-JUL-2000;	2000US-00620312.			
PR	03-AUG-2000;	2000US-00653450.			
PR	14-SEP-2000;	2000US-00662191.			
PR	19-OCT-2000;	2000US-00693036.			
PR	29-NOV-2000;	2000US-00727344.			
XX	(HYSE-) HYSEQ INC.				
PA	Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;				
XX	Tang YT, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;				
PI	Wang J, Zhou P, Goodrich R, Drmanac RT;				
PI	WPI; 2001-442253/47.				
XX	P-PSDB; AAM39588.				
DR	Novel nucleic acids and polypeptides, useful for treating disorders such				
XX	as central nervous system injuries.				
PT	Claim 1; SEQ ID NO 947; 10078pp; English.				
XX	The invention relates to human nucleic acids (AAI57798-AAI61369) and the				
CC	encoded polypeptides (AAM38642-AAM42213) with nootropic,				
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful				
CC	in gene therapy. A composition containing a polypeptide or polynucleotide				
CC	of the invention may be used to treat diseases of the peripheral nervous				
CC	system, such as peripheral nervous injuries, peripheral neuropathy and				
CC	localised neuropathies and central nervous system diseases, such as				
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic				
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the				
CC	utilisation of the activities such as: Immune system suppression,				
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic				
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,				
CC	assays for receptor activity, arthritis and inflammation, leukaemias and				
CC	C.N.S disorders. Note: The sequence data for this patent did not form				
CC	part of the printed specification				
XX	Sequence 659 BP; 197 A; 137 C; 173 G; 152 T; 0 U; 0 Other;				
SQ	Query Match 95.0%; Score 483.4; DB 4; Length 659;				
	Best Local Similarity 97.1%; Pred. No. 2.1e-126;				
	Matches 508; Conservative 0; Mismatches 1; Indels 14; Gaps 1;				
QY	1	GAGTTGTGAGGTGTGAGGGTGGCGTCTCTGCTGTCTGAGCTTTTCTGTCCCACTGAGA	60		
Db	135	GAGTTGTGAGGTGTGAGGGTGGCGTCTCTGCTGTCTGAGCTTTTCTGTCCCACTGAGA	194		
QY	61	CGCAGCT-----GTGTGAAATATATGATTTGGCGAGGAAGATCAACATATAGG	106		
Db	195	CGCAGCTGTATTCTGTTTGCAGTGTGAAATATGATTTGGCGAGGAAGATCAACATATAGG	254		
QY	107	CCTAGGCCGAGGAGAGTGTACCACTCTGAGCTGATTTGGGCTATGCTGGAGCCGGT	156		
Db	255	CCTAGGCCGAGGAGAGTGTACCACTCTGAGCTGATTTGGGCTATGCTGGAGCCGGT	314		
QY	167	GATGAGGAGCTCAGCAAGAGGAACCAACCACTGAAAGTCGGATCCTGCACCTGGTCAG	226		
Db	315	GATGAGGAGCTCAGCAAGAGGAACCAACCACTGAAAGTCGGATCCTGCACCTGGTCAG	374		
QY	227	GAGAGAGAAGATCAGGCTGAGCTGAGCTCAAGTGCCTGACCTGGAGGCTGCTC	286		
Db	375	GAGAGAGAAGATCAGGCTGAGCTGAGCTCAAGTGCCTGACCTGGAGGCTGCTC	434		
QY	287	CAGGAGCTGTCTCAGTCAAAAGATGGGGGTGAATGTGGAATGCTCTGATGACCAAGGG	346		
Db	435	CAGGAGCTGTCTCAGTCAAAAGATGGGGGTGAATGTGGAATGCTCTGATGACCAAGGG	494		
QY	347	AAGATTCTGCCAAATCAGAACCAATTTAAATGCCAGAGAGGTGACGGCAACCAAG	406		
Db	495	AAGATTCTGCCAAATCAGAACCAATTTAAATGCCAGAGAGGTGACGGCAACCAAG	554		
QY	407	TTTTAAATGAGCAAGCTGAACCAACCAAACTGTTTTTATCTAAGATATTGACTTA	466		
Db	555	TTTTAAATGAGCAAGCTGAACCAACCAAACTGTTTTTATCTAAGATATTGACTTA	614		
QY	467	AAAATATCAAAATAAACTTTTGCAGCTTTCTCCAAAAA	509		

QY 429 ACAACACAAAAGCTGTTTATCTAAGATATTGACTTAAATAATATCAAAATAAATCTTTG 488
 |||||
 DB 412 ACAACCCCAAAAGCTGTTTATTAAGATATTGACTTAAATAATATCGAAATAAATCTTTG 471
 |||||
 QY 489 CAGCTTTCTCC 499
 |||||
 DB 472 CAGCTTTCTCC 482

RESULT 9

AA569486
 ID AA569486 standard; cDNA; 661 BP.

AC AA569486;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #5290.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

FN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US008631.

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-839362/73.

DR P-PSDB; ABG05299.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX Claim 1; SEQ ID NO 5290; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AA564197-AA594564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from Wipo at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 661 BP; 212 A; 129 C; 158 G; 162 T; 0 U; 0 Other;

Query Match 77.4%; Score 393.8; DB 5; Length 661;

Best Local Similarity 98.7%; Pred. No. 4.7e-101;

Matches 439; Conservative 0; Mismatches 2; Indels 4; Gaps 4;
 QY 65 GCTGTGTGAAATATGATTTGGCGAGGAAGATCAACATATAGCCTAGGCCGAGGAGAACT 124
 |||||
 DB 217 GCAGTGTGAAATATGATTTGGCGAGGAAGATCAACATATAGCCTAGGCCGAGGAGAACT 276
 |||||
 QY 125 GTACCACTCTCCTGAGCTGATTTGGGCTATGCTGGAGCCCGGTGATGAGGAGCCTCAGCAA 184
 |||||
 DB 277 GTACCACTCTCCTGAGCTGATTTGGGCTATGCTGGAGCCCGGTGATGAGGAGCCTCAGCAA 336
 |||||
 QY 185 GAGGAACCAACCACTGAAAGTCGGGATCCTGCACCT-GGTGAGGAGAGAGAGATCA 243
 |||||
 DB 337 GAGGAACCAACCACTGAAAGTCGGGATCCTGCACCTGGGTGAGGAGAGAGAGATCA 396
 |||||
 QY 244 GGGTGCAGCTGAGACTCAAGTGCCTGACCTGGAA-GCTGATCTCCAGGAGCTGTCTCAGT 302
 |||||
 DB 397 GGGTGCAGCTGAGACTCAAGTGCCTGACCTGGAA-GCTGATCTCCAGGAGCTGTCTCAGT 456
 |||||
 QY 303 CAAAGACTGGGGTGAAATGTGAAATGGTCCTGATGACCCAGGGGAGATTTGCCAAAT 362
 |||||
 DB 457 CAAAGACTGGGGTGAAATGTGAAATGGTCCTGATGACCCAGGGGAGATTTGCCAAAT 516
 |||||
 QY 363 CAGAACAA-TTTTAAATGCCAGAA-GGAGTGCAGGCAACACACAGTTTTAAATGAAGAC 420
 |||||
 DB 517 CAGAACAA-TTTTAAATGCCAGGAGGAGTGCAGGCAACACACAGTTTTAAATGAAGAC 576
 |||||
 QY 421 AAGCTGAAACAAACACAAAACCTGTTTTTATCTAAGATATTGACTTAAAAATATCAAAATA 480
 |||||
 DB 577 AAGCTGAAACAAACACAAAACCTGTTTTTATCTAAGATATTGACTTAAAAATATCGAAATA 636
 |||||
 QY 481 AACTTTTGCAGCTTCTCCAAAAA 505
 |||||
 DB 637 AACTTTTGCAGCTTCTCCAAAAA 661

RESULT 10

ABA93876

ID ABA93876 standard; cDNA; 642 BP.

AC ABA93876;

DT 07-MAY-2002 (first entry)

DE Human G protein-coupled receptor NOV4 encoding cDNA SEQ ID NO:9.

KW Human; G protein-coupled receptor; receptor; GPCR; NOVX; cardiant;
 antiarteriosclerotic; cytosolic; immunomodulatory; anorectic; antiviral;
 antiinflammatory; coagulant; gene therapy; cardiomyopathy; hypertension;
 atherosclerosis; valve disease; obesity; prostate cancer; adenocarcinoma;
 uterus cancer; fertility; haemophilia; immunodeficiency; AIDS;
 graft versus host disease; Crohn's disease; multiple sclerosis; ss.

OS Homo sapiens.

EH Key Location/Qualifiers

CDS 207..542

FT /*tag= a

FT /product= "NOV4 protein"

XX WO200200691-A2.

PD 03-JAN-2002.

XX 27-JUN-2001; 2001WO-US020510.

PR 27-JUN-2000; 2000US-0214759P.

PR 31-OCT-2000; 2000US-0244546P.

PR 13-NOV-2000; 2000US-0248153P.

PR 11-JAN-2001; 2001US-0261014P.

XX 22-JAN-2001; 2001US-0263215P.

XX (CURA-) CURAGEN CORP.

PI Vernet CAM, Tchernev V, Putturajan M, Malyankar UM, Gusev V;
 PI Herrmann JL, Macdougall JR, Rastelli L, Zhong H, Spytek KA;
 PI Shenoy S, Gerlach VL, Gangolli EA, Stone DJ, Smithson G;
 DR WPI; 2002-090517/12.
 XX P-PSDB; ABB05753.

XX Nucleic acids encoding polypeptides, designated NOVX polypeptides, useful
 PT for treating a syndrome associated with a NOVX-associated disorder, e.g.
 PT cardiomyopathy, atherosclerosis, hypertension, valve diseases, obesity,
 PT and prostate cancer.

PS Claim 8; Page 43; 215pp; English.

XX The present invention describes novel human G protein-coupled receptors
 CC (GPCRs), designated NOVX proteins (i.e. a NOV protein where X is one of:
 CC 1a; 1b; 2; 3; 4; 5; 6a1; 6a2; 6a3; 6b; 6c; 6d; 7a; 7b; and 7c). NOVX
 CC proteins and polynucleotides have cardiac, antiarteriosclerotic,
 CC cytotatic, immunomodulatory, anorectic, antiviral, antiinflammatory and
 CC coagulant activities, and can be used in gene therapy. The NOVX proteins
 CC and polynucleotides are useful for treating cardiomyopathy,
 CC atherosclerosis, hypertension, valve diseases, obesity, prostate cancer,
 CC adenocarcinoma, uterus cancer, fertility, haemophilia, Crohn's disease,
 CC immunodeficiencies, AIDS, graft versus host disease or multiple
 CC sclerosis. The present sequence represents a human NOVX protein encoding
 CC cDNA sequence from the present invention

XX Sequence 642 BP; 197 A; 136 C; 160 G; 149 T; 0 U; 0 Other;

Query Match 66.3%; Score 337.4; DB 6; Length 642;

Best Local Similarity 80.2%; Pred. No. 4.2e-85;

Matches 437; Conservative 0; Mismatches 71; Indels 37; Gaps 2;

QY 1 GAGTTGTGAGGTGAGGGTCCGCTCTGCTGTGAGTCTTCTCTCCACTGAGA 60
 DB 95 GAGCTGTGAGGTGAGGGGCGACGTTCCAGCGCTGTGACTTTCTCTCTACTGAGA 154
 QY 61 CGCAGTGT-----GTGAAATATGATTG 84
 DB 155 CGCAGCTATAGTTCGCGAGCGCAGCTCTCCAGGAAGTGAATATGATTG 214
 QY 85 CGCAGGAGATCAACATATAGGCTAGGCGGAGGAGTGTACACCTCTGAGTGTAT 144
 DB 215 GCAGGAGAGATCAACATATAGGCTAGGCGGAGGAGTGTACACCTCTGAGTGTAT 274
 QY 145 TGGCCCTATCTGAGCGCCGCTGATGAGGAGCTTCAGCAAGAGAAACCACTCAAG 204
 DB 275 TGGGCTATCTTGAACCACTGATGAGAGCTTAAGAGAGAAACCACTCAAG 334
 QY 205 TCGGATCTGACCTGCTGAGGAGAGAGAGATCAGGTGCGAGCTGAGACTCAAT 264
 DB 335 TCGGATCTGACCTGATCAGAGAGAGAGAGATCAGGTGCGAGCTGAGACTCAAT 394
 QY 265 GCCTGACCTGAGAGCTGATCTCAGAGAGCTGTCTCAGTCAAGACTGGGGTGAATGCG 324
 DB 395 GCCTGACCTGAGAGCTGATCTCAGAGAGCTGTCTCAGTCAAGACTGGGGTGAATGCG 454
 QY 325 AATGTCCTGATGACAGGGGAGATCTGCGAAATCAGAACATTAATGCGAGA 384
 DB 455 AGTGGTACTGATGTCAGGGGAGATCTTACCAAGAGCAGACACTTAAATGCGAGA 514
 QY 385 AGGAGGTGACAGGCAACAGAGTCTTAAATGAGAGCAAGCTGAACCAACAACTGTT 444
 DB 515 AGCAGGTGAGGGGAATCAACAGTTTAAAGGAGATAGCTGAACCAACAACTGTT 573
 QY 445 TTTATCTAAGATATTTGACTTAAATATCAAAATAACTTTTTCAGCTTTTCCAAAA 504
 DB 574 TTTATCTAAGATATTTTACITTTAAATATCTTAAATGATTTTAAAGTTTCTCCAAAA 633
 QY 505 AAAAA 509
 DB 634 AAAAA 638

RESULT 11

AAK51918
 ID AAK51918 standard; cDNA; 673 BP.

XX AAK51918;

AC AAK51918;

XX 06-NOV-2001 (first entry)

XX Human polynucleotide SEQ ID NO 463.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.

OS Homo sapiens.

PN WO200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US004098.

XX 03-FEB-2000; 2000US-0046914.

PR 27-APR-2000; 2000US-00560875.

PR 20-JUN-2000; 2000US-00598075.

PR 19-JUL-2000; 2000US-00620325.

PR 01-SEP-2000; 2000US-00654936.

PR 15-SEP-2000; 2000US-00663561.

PR 20-OCT-2000; 2000US-00693325.

PR 30-NOV-2000; 2000US-00728422.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

PI Ma Y, Zhao QA, Wang D, Wang J, Ren F, Chen R, Wang ZW,

PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX WPI; 2001-476283/51.

XX P-PSDB; AAM78785.

XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.

PS Claim 1; Page 1680-1681; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication

XX Sequence 673 BP; 203 A; 145 C; 170 G; 155 T; 0 U; 0 Other;

Query Match 66.3%; Score 337.4; DB 4; Length 673;

Best Local Similarity 80.2%; Pred. No. 4.3e-85;

Matches 437; Conservative 0; Mismatches 71; Indels 37; Gaps 2;

QY 1 GAGTTGTGAGGTGAGGGTCCGCTCTGCTGTGAGCTTTCTGTCCTCCACTGAGA 60

DB 128 GAGCTGTGAGGTGAGGGGCGACGTTCCAGCGCTGTGAGCTTTCTCTCTACTGAGA 187

QY 61 CGCAGTCT-----GTGAAATATGATTG 84

KW Parkinson's disease; infertility; autoimmune disease; arthritis;
KW multiple sclerosis; allergy; wound healing; cytostatic; neurotropic;
KW immunosuppressive; neuroprotective; vulnery; hepatotropic; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

XX 5'UTR 1..173

XX CDS /*tag= a

XX 174..521

XX /*tag= b

XX /*product= "Human NOV4 protein"

XX 3'UTR 522..611

XX /*tag= c

XX WO200161009-A2.

XX 23-AUG-2001.

XX 15-FEB-2001; 2001WO-US004828.

XX 15-FEB-2000; 2000US-0182723P.

XX 15-FEB-2000; 2000US-0182724P.

XX 15-FEB-2000; 2000US-0182733P.

XX 22-FEB-2000; 2000US-0183896P.

XX 23-FEB-2000; 2000US-0184275P.

XX 23-FEB-2000; 2000US-0184482P.

XX 23-FEB-2000; 2000US-0184497P.

XX 24-FEB-2000; 2000US-0184744P.

XX 13-APR-2000; 2000US-0197083P.

XX 10-AUG-2000; 2000US-0224157P.

XX 18-SEP-2000; 2000US-0233405P.

XX 27-SEP-2000; 2000US-0236060P.

XX 02-JAN-2001; 2001US-0259414P.

XX 18-JAN-2001; 2001US-0262454P.

XX 14-FEB-2001; 2001US-00783429.

XX (CURA-) CURAGEN CORP.

XX Malyankar UM, Tchernev VT, Padigar M, Taupier RJ, Spytek KA;

XX Majumder K, Guo X, Spaderna SK, Boldog FL;

XX WPI; 2001-514775/56.

XX P-PSDB; AAE08583.

PT Isolated novel polypeptides useful for diagnosis of and treating cancer,
PT infertility, autoimmune diseases, arthritis, multiple sclerosis,
PT allergies, wound healing and hepatic disorders.

XX Claim 9; Page 14; 140pp; English.

XX The present sequence is a human NOV4 DNA. The NOVX protein has homology
XX with one of G-antigen (GAGE)-like protein, interferon, G-protein coupled
XX receptor (GPCR), hepatocyte nuclear factor or mast cell protease. The
XX NOVX is useful for treating or preventing a pathology associated with
XX NOVX. It is also useful for determining the presence or amount of NOVX
XX DNA in a sample, for identifying a potential therapeutic agent and in
XX gene therapy. It is also useful for determining the presence of or
XX predisposition to a disease associated with altered levels of NOVX. It is
XX also useful for the diagnosis and treatment of proliferative disorders,
XX e.g., cancer, immune disorders, hepatic disorders, e.g., cirrhosis, viral
XX infections, e.g., hepatitis, neurotoxicity, system-related disorders,
XX neurological disorders, e.g., Parkinson's disease, infertility,
XX autoimmune diseases, arthritis, multiple sclerosis, allergies and wound
XX healing

XX Sequence 611 BP; 196 A; 127 C; 152 G; 135 T; 0 U; 1 Other;

XX Query Match 63.4%; Score 322.8; DB 4; Length 611;

XX Best Local Similarity 79.5%; Pred No. 5, 5e-81;

XX Matches 435; Conservative 0; Mismatches 73; Indels 39; Gaps 3;

XX 1 GAGTTGTGAGGGTGTGAGGGTCCGTCCTGCTGTGAGACTTTTCTGCTCCACTGAGA 60

Db 62 GAGCTGTGAGAGTGTGAGGGACACGTTCCAGCGTCTGGACTTTCTCTCTACTGAGA 121
Qy 61 CGCAGCTGT-----GTGAAATATGATTG 84
Db 122 CGCAGCTATAGTTCGCGAGCCAGTCTCCAGGAACCTGAAATATGATTG 181
Qy 85 GCGAGGAGATCAACATATAGGCTAGGCGGAGGAGAGTGTACCACTCTCTCAGCTGAT 144
Db 182 GCGAGGAGATCAACATATAGGCTAGGCGGAGGAGAGTGTACCACTCTCTCAGCTGAT 241
Qy 145 TGGGCTATGCTGAGGCCGCTGATGAGGAGCTCTCAGCAAGAGGAAACCACTGAAAG 204
Db 242 TGGGCTATGCTTGAACCCACTGATGAGAGGCTTAAAGAGAGAAACCACTGAAAG 301
Qy 205 TCGGATCTTCACCTG--GTCAGGAGAGAGAGATCAGGTCAGGTCAGCTGAGACTCAA 262
Db 302 TCGGAATCTTACCTGACTCAGAGAGAGAGATGATCAGGTCAGCTGAGACTCAA 361
Qy 263 GTGCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGAGCTGGGGTGAATGT 322
Db 362 GTGCTGACCTGGAAGCGATCTCCAGGAGCTGTCTCAGCAAAAGCTGGGATGATGT 421
Qy 323 GGAATGCTCTGATGACCGAGGGAAGATTTCTCCAAATCAGACAAATTTAAATGCCA 382
Db 422 GAAGTGTGCTACTGATGTCAAGGGGAGATTTCTACCAAGGAGAGACTTTAAATGCCA 481
Qy 383 GAAGGAGGTGACAGGCAACCAAGCTTTAAATGAAGCAAGCTGAAACACACAAAACCTG 442
Db 482 GAAGCAGGTGAAGGGAATCACAGGTTTAAAGGAAGATAGCTGAAACACAC-AAAACCTG 540
Qy 443 TTTTATCTAAGATATTGACTTAAATATCAAAATTAATTTTTCAGCTTTCTCCAAA 502
Db 541 TTTTATATTAGATATTACTTTTAAATATCTTAAATAGTTTAAAGCTTTCTCCAAA 600
Qy 503 AAAAAA 509
Db 601 AAAAAA 607

RESULT 14

AAF59637

ID AAF59637 standard; cDNA; 580 BP.

XX AC AAF59637;

XX XX 24-APR-2001 (first entry)

XX DE Human cell cycle and proliferation protein CCVPR-48 cDNA, SEQ ID NO:102.

XX KW Cell cycle and proliferation protein; CCVPR; human; agonist; antagonist;
XX gene therapy; detection; gene therapy; transgenic animal disease model;
XX immune disorder; developmental disorder; cell signalling disorder;
XX cell proliferative disorder; cancer; tumour; anaemia; epilepsy;
XX arteriosclerosis; asthma; allergy; diabetes mellitus;
XX menstrual cycle disorder; bacterial infection; ss.

XX OS Homo sapiens.

XX XX WO200107471-A2.

XX PD 01-FEB-2001.

XX PF 21-JUL-2000; 2000WO-US019948.

XX PR 21-JUL-1999; 99US-0145075P.

XX PR 08-SEP-1999; 99US-0153129P.

XX PR 10-NOV-1999; 99US-0164647P.

XX (INCY-) INCYTE GENOMICS INC.

XX Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O;

XX Azimzai Y, Yang J, Lu DAM, Baughn MR, Patterson C, Shah P;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: April 24, 2004, 23:54:27 ; Search time 1340.3 Seconds
(without alignments)
16460.143 Million cell updates/sec

Title: US-10-051-835-11
Perfect score: 509
Sequence: 1 gagggtgaggtgtgaggg.....agctttctccaaaaaaa 509

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: gb_ba.*	BC062680 Homo sapi
2: gb_hgt.*	AR339143 Sequence
3: gb_in.*	AX921769 Sequence
4: gb_om.*	AX921771 Sequence
5: gb_ov.*	AJ318881 Homo sapi
6: gb_pat.*	AJ318880 Homo sapi
7: gb_ph.*	AJ318881 Homo sapi
8: gb_pl.*	BC009232 Homo sapi
9: gb_pr.*	AX359705 Sequence
10: gb_ro.*	AX226501 Sequence
11: gb_sts.*	268274 Human DNA s
12: gb_sy.*	AL117391 Human DNA
13: gb_un.*	AL096838 Human DNA
14: gb_vi.*	AL645949 Human DNA
15: em_ba.*	AX078298 Sequence
16: em_fun.*	BC009538 Homo sapi
17: em_hum.*	AF251237 Homo sapi
18: em_in.*	AX455517 Sequence
19: em_mu.*	AJ318878 Homo sapi
20: em_on.*	AR414108 Sequence
21: em_or.*	BD109661 EST and e
22: em_ov.*	AR273067 Sequence
23: em_pat.*	AR276648 Sequence
24: em_ph.*	AR406923 Sequence
25: em_pi.*	AX063181 Sequence
26: em_ro.*	AX368098 Sequence
27: em_sts.*	AR272357 Sequence
28: em_un.*	AR275938 Sequence
29: em_vi.*	AR406213 Sequence
30: em_hgt_hum.*	AX062442 Sequence
31: em_hgt_inv.*	AX367359 Sequence
32: em_hgt_other.*	AJ290447 Homo sapi
33: em_hgt_mus.*	AX370580 Sequence
34: em_hgt_pln.*	AF055473 Homo sapi
35: em_hgt_rod.*	BC018052 Homo sapi
36: em_hgt_mam.*	BC018052 Sequence
37: em_hgt_vrt.*	BD231797 Isolated
38: em_sy.*	I55851 Sequence 14
39: em_hgo_hum.*	BD132467 Isolated,
40: em_hgo_mus.*	AX285022 Sequence
41: em_hgo_other.*	U19143 Human GAGE-

ALIGNMENTS

RESULT 1
BC062680 548 bp mRNA linear PRI 11-DEC-2003
LOCUS BC062680 Homo sapiens cDNA clone MGC:71925 IMAGE:4619720, complete cds.
DEFINITION BC062680
ACCESSION BC062680.1 GI:38541861
VERSION BC062680.1
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 548)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.F., Schaefer,C.F., Bhat,N.K.,

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```

Location/Qualifiers
1. 548
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:71925 IMAGE:4619720"
/tissue_type="Placenta"
/clone_lib="NIH MGC 79"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
96. 431
/codon_start=1
/product="Unknown (protein for MGC:71925)"
/protein_id="AAH62680.1"
/db_xref="GI:138541862"
/translations="MTWRGRSTYRPRRSVPPPELIGFMLEPGDEEPQOEEPTPTSR
DPAPGGEREDOGAAETQVDLEADQLBLSQSKTGGCGNGPDDQGIKLPKSGQFKMP
EGGRDPQV"
156. 428
/note="GAGE; Region: GAGE protein. This family consists of
several GAGE and XAGE proteins which are found exclusively
in humans. The function of this family is unknown although
they have been implicated in human cancers"
/db_xref="CDD:pfam05831"

```

```

RESULT 2
LOCUS AR339143 linear 659 bp DNA
DEFINITION Sequence 634 from patent US 6569662.
ACCESSION AR339143
VERSION AR339143.1 GI:33726000
KEYWORDS
SOURCE
ORIGIN
REFERENCE
AUTHORS Tang, Y. T., Zhou, P. and Drmanac, R. T.
TITLE Nucleic acids and polypeptides
JOURNAL Patent: US 6569662-A 634 27-MAY-2003;
FEATURES
source
location/Qualifiers
1. .659
/organism="unknown"
/mol_type="genomic DNA"

Query Match 95.0%; Score 483.4; DB 6; Length 659;
Best Local Similarity 97.1%; Pred. No. 5.5e-113;
Matches 508; Conservative 0; Mismatches 1; Indels 14; Gaps 1;

QY 1 GAGTGTGAGGGGTGAGGTCGCGCTTCCTGCTGTCTGTCGACATTTTCTGTGCCACATGAGA 60
135 GAGTGTGAGGGGTGAGGTCGCGCTTCCTGCTGTCTGTCGACATTTTCTGTGCCACATGAGA 194

Db 61 CGCAGCT-----GTGTGAATATGATTTGGCGAGGAAGATCAACATATAGG 106
QY

```

195 CGACGCTGATTTCTGTTGCGAGTGGAATATGATTTGGCGAGGAGATCAACATATAGG 254
Db
107 CCTAGGCCGAGGAGAGTGTACACCTCTGAGCTGATTTGGGCTATGCTGGAGCCCGGT 166
Qy
255 CCTAGGCCGAGGAGAGTGTACACCTCTGAGCTGATTTGGGCTATGCTGGAGCCCGGT 314
Db
167 GATGAGGAGCTTCAGCAAGAGGAAACCAACAACTGAAAGTCCGGATCCTGCACTTGGTCAG 226
Qy
315 GATGAGGAGCTTCAGCAAGAGGAAACCAACAACTGAAAGTCCGGATCCTGCACTTGGTCAG 374
Db
227 GAGAGAGAGAGATCAGGGTGCAGCTGAGACTCAAGTCCCTGACCTGGAGAGCTGATCTC 286
Qy
375 GAGAGAGAGAGATCAGGGTGCAGCTGAGACTCAAGTCCCTGACCTGGAGAGCTGATCTC 434
Db
287 CAGGAGCTGTCTCAGTCAAGAGCTGGGGTGAATGTGGAATGGTCTGTGATGACCCAGGG 346
Qy
435 CAGGAGCTGTCTCAGTCAAGAGCTGGGGTGAATGTGGAATGGTCTGTGATGACCCAGGG 494
Db
347 AAGATTTCTGCCAAATCAGAACTTTAAATGCGAAGAGAGTGCAGGCAACCCAG 406
Qy
495 AAGATTTCTGCCAAATCAGAACTTTAAATGCGAAGAGAGTGCAGGCAACCCAG 554
Db
407 GTTTAAATGAAGACAGCTGAAACCAACAACTGTTTATCTAAGATATTTGACTTA 466
Qy
555 GTTTAAATGAAGACAGCTGAAACCAACAACTGTTTATCTAAGATATTTGACTTA 614
Db
467 AAAATATCAAAATAAATTTTTCAGCTTTCTCCAAAAA 509
Qy
615 AAAATATCAAAATAAATTTTTCAGCTTTCTCCAAAAA 657
Db

RESULT 3
AX921769 550 bp DNA linear PAT 18-DEC-2003
LOCUS
DEFINITION
Sequence 109 from Patent WO02068649.
AX921769
ACCESSION
AX921769.1 GI:40215336
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
1

REFERENCE
AUTHORS
JOURNAL
FEATURES
source
1. .550
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
Patent: WO 02068649-A 109 06-SEP-2002;
Curagen Corporation (US)
Location/Qualifiers
ORIGIN
Query Match 89.9%; Score 457.8; DB 6; Length 550;
Best Local Similarity 96.8%; Pred. No. 2e-106;
Matches 483; Conservative 0; Mismatches 2; Indels 14; Gaps 1;
Qy 25 GTTCCTGCTGCTGGAGTCTTCTGCTCCCACTGAGAGCGAGCT-----GTG 70
Db 1 GTTCCTGCTGCTGGAGTCTTCTGCTCCCACTGAGAGCGAGCTGTTCTGTTGCAAGT 60
Qy 71 TGAATATGATTTGGCGAGGAGATCAATATAGGCTTAGGCGGAGGAGTGTACCA 130
Db 61 TGAATATGATTTGGCGAGGAGATCAATATAGGCTTAGGCGGAGGAGTGTACCA 120
Qy 131 CCTCCTGAGTGTGGGCTATGCTGGAGCCCGGTGATGAGAGCTCAGCAAGAGGAA 190
Db 121 CCTCCTGAGTGTGGGCTATGCTGGAGCCCGGTGATGAGAGCTCAGCAAGAGGAA 180
Qy 191 CCACCAACTGAAAGTCCGGATCTGCTGAGCCCGGTGATGAGAGCTCAGGCTGCA 250
Db 181 CCACCAACTGAAAGTCCGGATCTGCTGAGCCCGGTGATGAGAGCTCAGGCTGCA 240
Qy 251 GCTGAGACTCAAGTCCCTGAGCTGGAAGCTGATCTCCAGGAGCTGCTCAGTCAAGACT 310
Db 241 GCTGAGACTCAAGTCCCTGAGCTGGAAGCTGATCTCCAGGAGCTGCTCAGTCAAGACT 300
Qy 311 GGGGCTGATGTGGAATGGTCTGATGACCCAGGGGAGATTTGCCCCAAATCAGACAA 370
Db 301 GGGGCTGATGTGGAATGGTCTGATGACCCAGGGGAGATTTGCCCCAAATCAGACAA 360
Qy 371 TTTAAATGCCAGAGGAGGTGCAGGCAACCAAGTTTAAATGAAGACAAGCTGAAAC 430
Db 421 TTTAAATGCCAGAGGAGGTGCAGGCAACCAAGTTTAAATGAAGACAAGCTGAAAC 420

251 GCTGAGACTCAAGTCCCTGAGCTGGAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGACT 310
Db
241 GCTGAGACTCAAGTCCCTGAGCTGGAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGACT 300
Qy
311 GGGGCTGATGTGGAATGGTCTGATGACCCAGGGGAGATTTGCCCCAAATCAGACAA 370
Db
301 GGGGCTGATGTGGAATGGTCTGATGACCCAGGGGAGATTTGCCCCAAATCAGACAA 360
Qy
371 TTTAAATGCCAGAGGAGGTGCAGGCAACCAAGTTTAAATGAAGACAAGCTGAAAC 430
Db
421 TTTAAATGCCAGAGGAGGTGCAGGCAACCAAGTTTAAATGAAGACAAGCTGAAAC 420
Qy
431 AACACAAAATGTTTTTATCTAAGATATTTGACTTAAATATCAAAATCAAACTTTTGA 490
Db
421 AACACAAAATGTTTTTATCTAAGATATTTGACTTAAATATCAAAATCAAACTTTTGA 480
Qy
491 GCTTTCTCCAAAAA 509
Db
481 GCTTTCTCCAAAAA 499
Qy

RESULT 4
AX921771 494 bp DNA linear PAT 18-DEC-2003
LOCUS
DEFINITION
Sequence 111 from Patent WO02068649.
AX921771
ACCESSION
AX921771.1 GI:40215337
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
1

REFERENCE
AUTHORS
JOURNAL
FEATURES
source
1. .494
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
Patent: WO 02068649-A 111 06-SEP-2002;
Curagen Corporation (US)
Location/Qualifiers
ORIGIN
Query Match 89.6%; Score 456; DB 6; Length 494;
Best Local Similarity 97.2%; Pred. No. 5.7e-106;
Matches 480; Conservative 0; Mismatches 0; Indels 14; Gaps 1;
Qy 25 GTTCCTGCTGCTGGAGTCTTCTGCTCCCACTGAGAGCGAGCT-----GTG 70
Db 1 GTTCCTGCTGCTGGAGTCTTCTGCTCCCACTGAGAGCGAGCTGTTCTGTTGCAAGT 60
Qy 71 TGAATATGATTTGGCGAGGAGATCAATATAGGCTTAGGCGGAGGAGTGTACCA 130
Db 61 TGAATATGATTTGGCGAGGAGATCAATATAGGCTTAGGCGGAGGAGTGTACCA 120
Qy 131 CCTCCTGAGTGTGGGCTATGCTGGAGCCCGGTGATGAGAGCTCAGCAAGAGGAA 190
Db 121 CCTCCTGAGTGTGGGCTATGCTGGAGCCCGGTGATGAGAGCTCAGCAAGAGGAA 180
Qy 191 CCACCAACTGAAAGTCCGGATCTGCTGAGCCCGGTGATGAGAGCTCAGGCTGCA 250
Db 181 CCACCAACTGAAAGTCCGGATCTGCTGAGCCCGGTGATGAGAGCTCAGGCTGCA 240
Qy 251 GCTGAGACTCAAGTCCCTGAGCTGGAAGCTGATCTCCAGGAGCTGCTCAGTCAAGACT 310
Db 241 GCTGAGACTCAAGTCCCTGAGCTGGAAGCTGATCTCCAGGAGCTGCTCAGTCAAGACT 300
Qy 311 GGGGCTGATGTGGAATGGTCTGATGACCCAGGGGAGATTTGCCCCAAATCAGACAA 370
Db 301 GGGGCTGATGTGGAATGGTCTGATGACCCAGGGGAGATTTGCCCCAAATCAGACAA 360
Qy 371 TTTAAATGCCAGAGGAGGTGCAGGCAACCAAGTTTAAATGAAGACAAGCTGAAAC 430
Db 421 TTTAAATGCCAGAGGAGGTGCAGGCAACCAAGTTTAAATGAAGACAAGCTGAAAC 420

Db 361 TTTAAATGCCAGAGGCTGACAGGCAACACAGGTTTAAATGAAGACAGCTGAAC 420

Qy 431 AACACAAACTCTTTTATCTAGATATTTGACTTAAATAATCAAAATAAATTTGCA 490

Db 421 AACACAAACTCTTTTATCTAGATATTTGACTTAAATAATCAAAATAAATTTGCA 480

Qy 491 GCTTTCTCCAAAAA 504

Db 481 GCTTTCTCCAAAAA 494

RESULT 5

HS3118881

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. .493

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="X"

/map="Xp11.21-Xp11.22"

/tissue_type="Ewing sarcoma"

1. .493

/gene="XAGE-3"

61. .396

/gene="XAGE-3"

/codon_start=1

/product="XAGE-3 protein"

/protein_id="CAC83008.1"

/db_xref="GI:18157212"

/db_xref="SWISS-PROT:O8WTP9"

/translation="MIWRGSTYRPRRSPVPPPELIGPMLPDEPQEEPPPTESR

DPAPQQRREDQAAETQVPLEADLQELSQSKTGCGCGNPDQDQKILPKSEQFMP

EGGDRPQPV"

ORIGIN

Query Match 86.3%; Score 439.4; DB 9; Length 493;

Best Local Similarity 99.8%; Pred. No. 1e-101;

Matches 440; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 69 TGTGAATATGATTTGGCAGGAGATCAACATATAGCCCTAGCCGAGGAGAGTGTAC 128

Db 53 TGTGAATATGATTTGGCAGGAGATCAACATATAGCCCTAGCCGAGGAGAGTGTAC 112

Qy 129 CACCTCTGAGCTGATTTGGGCTTATGCTGGAGCCCGGTGATGAGGACCTCAGCAAGAG 188

Db 113 CACCTCTGAGCTGATTTGGGCTTATGCTGGAGCCCGGTGATGAGGACCTCAGCAAGAG 172

Qy 189 AACCAACCACTGAAGTCGGATCTCTGACCTGTCAGGAGAGAGAGATCAGGGTG 248

Db 173 AACCAACCACTGAAGTCGGATCTCTGACCTGTCAGGAGAGAGAGATCAGGGTG 232

Qy 249 CAGCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGA 308

Db 233 CAGCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGA 292

Qy 309 CTGGGGTGAATGTGAAATGGTCTGATGACGAGGGAAGATTCTGCCAAATCAGAAC 368

Db 293 CTGGGGTGAATGTGAAATGGTCTGATGACGAGGGAAGATTCTGCCAAATCAGAAC 352

Qy 369 AATTTAAATGCCAGAGAGGTGACAGGCAACCCAGAGTTTAAATGAAGACAGCTGAA 428

Db 353 AATTTAAATGCCAGAGAGGTGACAGGCAACCCAGAGTTTAAATGAAGACAGCTGAA 412

Qy 429 ACAACACAAACTCTTTTATCTAGATATTTGACTTAAATAATCAAAATAAATTTTG 488

Db 413 ACAACACAAACTCTTTTATCTAGATATTTGACTTAAATAATCAAAATAAATTTTG 472

Qy 489 CAGCTTTCTCCAAAAA 509

Db 473 CAGCTTTCTCCAAAAA 493

RESULT 6

HS3118880

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. .620

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="X"

/map="Xp11.21-Xp11.22"

/tissue_type="Ewing sarcoma"

1. .620

/gene="XAGE-2"

189. .524

/gene="XAGE-2"

/codon_start=1

/product="XAGE-2 protein"

/protein_id="CAC83007.1"

/db_xref="GI:18157210"

/db_xref="SWISS-PROT:Q96GT9"

/translation="MSWRGSTYRPRRSLQPPPELIGAMLEPDEPPEKPKPTKSR

NPTDKREDQGAAEIQVPLEADLQELCQTKTGDCGEGTDYKGIPLKAEHFMP

EAGEGKSVQ"

ORIGIN

Query Match 66.3%; Score 337.4; DB 9; Length 620;

Best Local Similarity 80.2%; Pred. No. 1.4e-75;

Matches 437; Conservative 0; Mismatches 71; Indels 37; Gaps 2;

Qy 1 GAGTTGTAGGCTGTGAGGTCGCGCTTCTGCTGCTGAGACTTTTCTGTCCCACTGAGA 60

Db 77 GAGCTGTGAGCTGTGAGGGGACGCTTCAGCCGCTCTGACTCTTTCTCTCTACTGAGA 136

Db 182 GCGAGGAAGATCAACATATAGGCTAGGCCAAGAAGAGCTTTACAGCCTCTGAGCTGAT 241

QY 145 TGGGCTATGCTGGAGCCGGTGTGATGAGGAGCTCAGCAAGAGGAGCAACCACTGAAG 204

Db 242 TGGGCTATGCTTGAACCACTGTGATGAGGAGCTTAAGNAGAGAAACCACTGAAG 301

QY 205 TGGGATCTGCACTG--GTGAGGAGAGAGAGATCAGGGTGCAGCTGAGACTCA 262

Db 302 TCGGAATCTCACCTGACTCNAGAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTCA 361

QY 263 GTGCTGACTGGAAGCTGATCTCAGGAGCTGTCTCAGTCAAGAGCTGGGGTGAATGT 322

Db 362 GTGCTGACTGGAAGCGATCTCAGGAGCTATGTGACAAAGACTGGGATGATGT 421

QY 323 GGAATGCTGATGATGACAGGGAGAGATTCGCCAAATCAGAAATTAATAATGCCA 382

Db 422 GAAGTGGTACTGATGTCAAGGGAGAGATTCACCAAGAGAGAGCACTTTAAATGCCA 481

QY 383 GAAGGAGGTGACAGCAACACAGCTTTAAATGAGACAGAGCTGAACACACAAACCTG 442

Db 482 GAAGCAGGTGAAGGGAATCAGAGGTTTAAAGAGAGATGAGCTGAACACACAACTG 540

QY 443 TTTTATCTAGATATTGCACTTAAATATCAAAATTAACCTTTTCGAGCTTTCCCAA 502

Db 541 TTTTATATTAGATATTTTACTTTAAATATCTTAATAAGTTTAAAGCTTTTCCTCAA 600

QY 503 AAAAAA 509

Db 601 AAAAAA 607

RESULT 10

HSL185E6A/c HSL185E6A 20587 bp DNA linear PRI 23-NOV-1999

LOCUS Human DNA sequence from cosmid L129H7, Huntington's Disease Region, chromosome 4p16.3 contains Pseudogene and CpG island. 268274

DEFINITION

ACCESSION 268274.1 GI:1130686

VERSION 4p16.3; CpG island; pseudogene.

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 20587)

AUTHORS Buck,D.

TITLE Direct Submission

JOURNAL Submitted (12-DEC-1995) Sanger Centre, Hinxton, Cambridgeshire, CB10 1RQ, England. E-mail enquiries: humquery@sanger.ac.uk

COMMENT

IMPORTANT: This sequence is not the entire insert of clone L185E6. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions. In addition this sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where unambiguous sequence has not been obtained, there is an annotation using the 'unsure' feature key. Where the size of a problem region is known only approximately the sequence contigs on either side will be submitted as separate entries, with annotation to indicate the order and positions of neighbours and what is known about the size and nature of the gap. Contigs have entryname suffixes (A, B, ...) indicating the order of the contigs with respect to the clone.

The sequence from clone L185E6 has been finished in more than one contig. This sequence (L185E6A) is separated from the following one (L185E6B) by a 500bp gap based on restriction data. 55bp VNTR made closure of gap impossible.

The true right end of clone L21F12 is at 10842 in this sequence. The true left end of clone L185E6 is at 1 in this sequence. L185E6 is from cosmid library LA04NC01 constructed at the Human Genome Center, Los Alamos National Laboratory, NM 87545 under the auspices of the U.S. Department of Energy. The library was constructed using

flow-sorted human chromosome 4 from a Hamster-Human hybrid cell line (UV20HL21-27) containing human chromosomes 4, 8 and 21.

VECTOR: sCos1

L185E6 is contained in a clone contig spanning 2Mb which is described in Baxendale et al, Nature Genetics 4 (1993) 181-186. See also Myers et al, Cytogenet Cell Genet. 66 (1994) 218-230.

FEATURES

source 1. 20587

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="4"

/map="4p16.3"

/clone="LA04NC01-185E6"

/cell_line="UV20HL21-27"

/clone_lib="LA04NC01"

/complement(1370..1739)

/note="match: STS L00885"

2380..4066

/notes="L1 element fragment"

4115..4400

/notes="Alu repeat: matches 1..308 of consensus"

4427..4736

/note="L1 element fragment"

4778..5099

/note="L1 element fragment"

5203..5333

/notes="MSTC element fragment"

5632..5721

/notes="L1 element fragment"

5980..6714

/notes="matches Trk-2h oncogene - 41aa of human RL7A_HUMAN M36072"

/pseudo

/codon_start=1

6747..6778

/notes="16 copies of 2 mer 100 % conserved"

10906..11082

/notes="Alu repeat: matches 1..191 of consensus"

11083..11214

/note="Alu repeat: matches 175..308 of consensus"

11352..11506

/notes="Alu repeat: matches 308..151 of consensus"

13249..13799

/note="putative CpG island"

15946..16238

/notes="Alu repeat: matches 1..308 of consensus"

17069..17143

/note="MIR element fragment"

17589..17689

/notes="Alu repeat: matches 308..210 of consensus"

17695..17739

/notes="Alu repeat: matches 170..114 of consensus"

18268..18369

/note="Alu repeat: matches 155..38 of consensus"

19668..19742

/note="MIR element fragment"

20313..20524

/note="4 copies of 53 mer 83 % conserved"

ORIGIN

Query Match 57.8%; Score 294.2; DB 9; Length 20587;

Best Local Similarity 82.5%; Pred. No. 2e-64;

Matches 387; Conservative 0; Mismatches 73; Indels 9; Gaps 4;

QY 1 GAGTTGAGGGGTGTGAGGGTCCGGTCTCTGCTGTGAGACTTTTCTGTGCCACTGAGA 60

18213 GACTGTGAGGGGTGTGAGGGTTCGGTTCCTGCTCTTACCCACTACTGAGA 18154

QY 61 CGCAGCTGTGTGAATATGATTTGGCAGGAAGATCAATATAGCCTAGGCCGAGGAG 120

18153 AGCAGCGGTGGAATATGATTTGGCAAGGAAGATCAGCATGTAGGCTTAGCAAGACG 18094

121 AAGTGTACCACTCTGAGCTGATGGCCCTATGCTGGAGCCCGGTGATGAGAGCCTCA 180
 Db GTATGTGACGCTCTGAGTACTAGGCTGCTGTGAGCCAGTGTGAGAGCCTCA 18034
 181 GCAAGAGAACCACTCACTGAGTGGGATCTGACCTGGTGTGAGAGAGAGAGAGAG 240
 Db GCAAGAGAGAACCACTCACTGAGTGGGATCTGACCTGGTGTGAGAGAGAGAGAGAG 17976
 241 TCAGGGTGCAGCTGAGACTCAAGTGCCTGACCTGGAGCTGATCTCCAGGAGCTGCTCA 300
 Db TCAGGGTGCCTGAGATCTGCTGTGACAGAGAGCTGATCTCCGAGAGCTGCTC- 17917
 301 GTCAAGAGCTGGGGTGAATGTGGAATGCTCTGATGACAGGGGAGAGATTTCTGCAAA 360
 Db -----AAGAGCTGGGATGAATGTGAGATGCTCTGATGCTC-GGGGAATATTCTGCCGAA 17862
 361 ATCAGAACCAATTTAAATGTCAGAGAGGAGTGCAGGACCAACACAGGTTTAAATGAAGAC 420
 Db ATCAGAGCAATTTAAATGTCAGAGAGGAGTGCAGGACCAACACAGGTTTCCAGGAAGAC 17802
 421 AAGCTGAACCAACACAACTGTTTTATCTAAGATATTGACTTAAAA 469
 Db AAGCTGAACCAATGC-AAAGCTGTTTTATATTAGATAGTACTTAAAA 17754

RESULT 11
 HS193G15/c
 LOCUS Human DNA sequence from clone RP1-193G15 on chromosome 4, complete
 DEFINITION sequence.
 ACCESSION AL117391
 VERSION AL117391.1 GI:5911814
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 62493)
 AUTHORS McMurray, A.
 TITLE Direct Submission
 JOURNAL Submitted (05-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk
 COMMENT Center: Wellcome Trust Sanger Institute
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Eni, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPep; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/C_elegans/wormpep RP1-193G15 is from the library RPCI-1 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm VECTOR: pCYPAC2.

FEATURES
 source Location/Qualifiers
 1..62493
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="4"
 /clone="RP1-193G15"
 /clone_lib="RPCI-1"
 ORIGIN
 Query Match 57.8%; Score 294.2; DB 9; Length 62493;
 Best Local Similarity 82.5%; Pred. No. 2.1e-64;
 Matches 387; Conservative 0; Mismatches 73; Indels 9; Gaps 4;
 QY 1 GAGTTGTGAGGGTGTGAGGGTTCGGTTCTCTGCTGTCTGAGACTTTTCTGCTCCACTCAGA 60
 Db GACCTGTGAGGGTGTGAGGGTTCGGTTCTCTGCTGTCTGAGACTTTTCTGAGACTTCTGAG 54245
 QY 61 CGAGCTGTGTGAATATGATTTGGCGAGGAAGATCAACATATAGGCTTAGGCGGAGGAG 120
 Db 54244 ASCAGCCGTTTGAATATGATTTGGCAAGGAAGATCAGCATGTAGGCTTAGACCAAGACG 54185
 QY 121 AAGTGTACCACTCTGAGCTGATTGGGCTTATGCTGGAGCCCGGTGATGAGAGCCTCA 180
 Db 54184 GTAATGTGAGCTCTGAGCTTACTGGGCTTGTGCTGAGCCAGTGTAGGAGCTCTCA 54125
 QY 181 GCAAGAGAACCACTCACTGAGTGGGATCTGCACTGCTGTCAGGAGAGAGAGAGAG 240
 Db 54124 GCAAGAGAACCACTCACTGAGTGGGATCTGCACTGCTGTCAGGAGAGAGAGAGAG 54067
 QY 241 TCAGGGTGCAGCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTGCTCA 300
 Db 54066 TCAGGGTGCCTGAGATTTCTGCTTGACAGGAGCTGATCTCCGAGAGCTGCTC- 54008
 QY 301 GTCAAGAGCTGGGGTGAATGTGGAATGCTCTGATGACAGGGGAGAGATTTGTCGCAA 360
 Db 54007 ----AAAGACTGGGATGAATGTGAGATGCTCTGATGCTC-GGGGAATATTCTGCCGAA 53953
 QY 361 ATCAGAACCAATTTAAATGTCAGAGAGGAGTGCAGGACCAACACAGGTTTAAATGAAGAC 420
 Db 53952 ATCAGAGCAATTTAAATGTCAGAGAGGAGTGCAGGACCAACACAGGTTTCCAGGAAGAC 53993
 QY 421 AAGCTGAACCAACAACTGTTTTATCTAAGATATTGACTTAAAA 469
 Db 53892 AAGCTGAACCAATGC-AAAGCTGTTTTATATTAGATAGTACTTAAAA 53845
 RESULT 12
 HSJ193G15/c
 LOCUS Human DNA sequence from clone 193G15 on chromosome 4 Contains the
 DEFINITION 5' part of the RGS12 gene for regulator of G-protein signalling 12,
 an RPL7A (60S Ribosomal Protein 7A (SURF3)) pseudogene, ESTs, an
 STS, GSSs and a putative CpG island, complete sequence.
 ACCESSION AL096838
 VERSION AL096838.1 GI:5459253
 KEYWORDS HTG; 60S Ribosomal Protein 7A; CpG Island; G-protein; RGS12; RPL7A;
 SURF3.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 62493)
 AUTHORS McMurray, A.
 TITLE Direct Submission
 JOURNAL Submitted (10-JUL-1999) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 COMMENT requests: clonerequest@sanger.ac.uk
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Eni, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPep; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/C_elegans/wormpep RP1-193G15 is from the library RPCI-1 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm VECTOR: pCYPAC2.

only a small overlap as described above.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
Em., EMSL; Sw., SWISSPROT; Tr., TREMBL
IMPORTANT: This sequence is not the entire insert of clone 193G15. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.

The true left end of clone 1B6 is at 6371 in this sequence. The true right end of clone 361H4 is at 104 in this sequence. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

This sequence was generated from part of bacterial clone contigs of human chromosome 4, constructed by the Sanger Centre Chromosome 4 Mapping Group.
193G15 is from the library RPC11 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/> VECTOR: pcYPAC2.

FEATURES

Location/Qualifiers
1..62493
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/clone="RP1-193G15"
/clone_lib="RPC1-1"

ORIGIN

Query Match 57.8%; Score 294.2; DB 9; Length 62493;
Best Local Similarity 82.5%; Pred. No. 2.1e-64;
Matches 387; Conservative 0; Mismatches 73; Indels 9; Gaps 4;
QY 1 GAGTTGTGAGGGGTGAGGGTCCGCTTCCTGCTGCTGGACTTTTCTGTCCCACTGAGA 60
Db 54304 GACCTGTGAGGGTGTGAGGGTTCGCTTCTGGTCTGGACTTTTACCACTACTGAGA 54245
QY 61 CGCAGCTGTGGAATATGATTTGGCGAGGAAGATCAACATATAGGCTAGCCGAGGAG 120
Db 54244 AGCAGCGGTTTGAAATATGATGTGGCAAGGAAGATCAGCATGTAGGCTTTAGACCAAGACG 54185
QY 121 AAGTGTACCACTCTGAGCTGATTTGGGCTATGCTGAGCCCGGTGATGAGGACCTCA 180
Db 54184 GTATGTGACGCTCTGAGCTAATCTGGGCTGTGCTTGAGCCCACTGATGAGCAGCTCA 54125
QY 181 GCAAGAGGAACCAACCACTGAAAGTCGGATCTCTGCACTGTGTACGAGGAGGAAGA 240
Db 54124 GCAAGAGGAACCAACCACTGAAAGTCGGGCTCTT--ACCTGGCCAGGAGGAGGAAGA 54067
QY 241 TCAGGTGACGCTGAGACTCAAGTCTGCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCA 300
Db 54066 TCAGGTGCTGTGAGATTTCTGTGCTTGACCAAGAGCTGATCTCCGGAGCTGTCTC- 54008
QY 301 GTCAAGACTGGGGGTGAATGTGGAATAGTCTGATGACACGGGGAAGATTTCTCCAAA 360
Db 54007 ----AAAGACTGGGATGAATGTGGAGATGGTCTGATGTCC--GGGGAATATTCTCCGAA 53953
QY 361 ATCAGAACATTTAAATCCAGAGGGGTGACAGGCAACCAAGGTTTAAATCAAGAC 420
Db 53952 ATCAGAGCAATTTAAATCCAGAGGGGTGAAGGCAACCAAGGTTTCCAGGAAGAC 53893
QY 421 AAGCTGAAACACACAAAAGTGTTTTATTAAGATATTGTACTTAAAA 469
Db 53892 AAGCTGAAACAATGC-AAAGTGTGTTTATTTAGATACGTGACTTAAAA 53845

RESULT 13
AL645949/c 145151 bp DNA linear PRI 23-OCT-2002
LOCUS
DEFINITION Human DNA sequence from clone RP11-357G3 on chromosome 4, complete

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AL645949 GI:24366471
HTG.
Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (16-OCT-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
Humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Oct 24, 2002 this sequence version replaced gi:24210295.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.: EMBL; Sw.: SWISSPROT; Tr.: TREMBL; Wp.: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP11-357G3 is from the library RPC1-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pBACE3.6. Location/Qualifiers

FEATURES

1..145151
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/clone="RP11-357G3"
/clone_lib="RPC1-11.2"

ORIGIN

Query Match 57.8%; Score 294.2; DB 9; Length 145151;
Best Local Similarity 82.5%; Pred. No. 2.2e-64;
Matches 387; Conservative 0; Mismatches 73; Indels 9; Gaps 4;
QY 1 GAGTTGTGAGGGGTGAGGGTCCGCTTCCTGCTGCTGGACTTTTCTGTCCCACTGAGA 60
Db 102337 GACCTGTGAGGGGTGAGGGTTCGCTTCTGCTGCTTACCCACTACTGAGA 102278
QY 61 CGCAGCTGTGTGAAATATGATTTGGCGAGGAAGATCAACATATAGGCTAGCCGAGGAG 120
Db 102277 AGCAGCGGTTTGAATATGATTTGGCAAGGAAGATCAGCATGTAGCTTAGACCAAGACG 102218
QY 121 AAGTGTACCACTCTGAGCTGATTTGGGCTTATGCTGAGCCCGGTGATGAGGAGCTCA 180
Db 102217 GTATGTGAGGCTCTGAGCTTAACTGGGCTGTGCTTGGAGCCAGTGATGAGGAGCTCA 102158
QY 181 GCAAGAGGAACCAACCACTGAAAGTGGGATCTGCACTGGTGTGAGGAGGAAGAAGA 240
Db 102157 GCAAGAGGAACCAACCACTGAAAGTGGGATCTT--ACCTGGCCAGGAGGAGGAGAGA 102100

/tissue_type="Lung, carcinoma, large cell

undifferentiated"

/clone_lib="NIH_MGC_69"

/lab_host="DH10B"

/note="Vector: PCMV-SPORT6"

1. .529

gene

/gene="GAGED2"

/note="synonyms: XAGE1, XAGE-1"

/db_xref="LocusID:9503"

/db_xref="MIM:300289"

184. .429

/codon_start=1

/product="GAGED2 protein"

/protein_id="AAH09538.1"

/db_xref="GI:16306935"

/db_xref="LocusID:9503"

/translation="MESPKNQOLKVLHLSRQKKIRIQLRSQCATWKVICKSCI

SQTFGINLDLGSVKVXLIIPKEHCKPEAGEEPQV"

ORIGIN

Query Match	42.1%	Score 214.4	DB 9	Length 529
Best Local Similarity	80.6%	Pred. No. 4.4e-44		
Matches 287	Conservative 0	Mismatches 66	Indels 3	Gaps 3
QY 155	CTGGAGCCCGGTGATGAGGAGCCCTCAGCAGAGGAAACCACTGAAAGTCGGGATCCT	214		
Db 171	CAGGAGCCCGGTGATGAGGAGCCCTCAGCAGAGGAAACCACTGAAAGTCGGGATCCT	230		
QY 215	GCACCTGGTCAGGAGAGAGAGATCAGGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC	274		
Db 231	ACACCTGGTCAGGAGAGAGAGATCAGGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC	290		
QY 275	GAAGCTGATCTCCAGGAGTGTCTCAGTCAAGAGTGGGGTGAATGTGGAAT-GGTCC	333		
Db 291	GAAGCTGATCTCCAGGAGTGTCTCAGTCAAGAGTGGGGTGAATGTGGAATGTGGAAT	350		
QY 334	TGATGACCAAGGGAAGATTCTGCCAAATTCAGAAATTCAGAAATTCAGAAATTCAGAA	393		
Db 351	CGGCTCAAGGTGAGATAATACCTAAGAGGAACACTGTAAATGCCAGAGCAGGTGA	410		
QY 394	CAGGCAACCAAGGTTTAAATGAGACAGTGAACCAACCAACCACTGTTTATCTAA	453		
Db 411	AGAGCAACCAAGGTTTAAATGAGACAGTGAACCAACCACTGTTTATCTAA	469		
QY 454	GATATTGACTTAAATATCAAAATAACTTTTCAGCTTTCTCCAAAAA	509		
Db 470	GATATTGACTTAAATATCAAAATAACTTTTCAGCTTTCTCCAAAAA	524		

Search completed: April 25, 2004, 18:57:27
Job time : 1351.3 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: January 17, 2004, 17:33:54 ; Search time 1687.16 Seconds
(without alignments)
16221.632 Million cell updates/sec

Title: US-10-051-835-1
Perfect score: 669
Sequence: 1 cacaacgagcagccagcatt.....tttttgcgcatgctgcatc 669

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :				GenEmbl:			
1:	gb	ba:		1:	gb	ba:	
2:	gb	htg:		2:	gb	htg:	
3:	gb	in:		3:	gb	in:	
4:	gb	om:		4:	gb	om:	
5:	gb	ov:		5:	gb	ov:	
6:	gb	pat:		6:	gb	pat:	
7:	gb	ph:		7:	gb	ph:	
8:	gb	pl:		8:	gb	pl:	
9:	gb	pr:		9:	gb	pr:	
10:	gb	ro:		10:	gb	ro:	
11:	gb	sts:		11:	gb	sts:	
12:	gb	sy:		12:	gb	sy:	
13:	gb	un:		13:	gb	un:	
14:	gb	vi:		14:	gb	vi:	
15:	em	ba:		15:	em	ba:	
16:	em	fun:		16:	em	fun:	
17:	em	hum:		17:	em	hum:	
18:	em	in:		18:	em	in:	
19:	em	mu:		19:	em	mu:	
20:	em	om:		20:	em	om:	
21:	em	or:		21:	em	or:	
22:	em	ov:		22:	em	ov:	
23:	em	pat:		23:	em	pat:	
24:	em	ph:		24:	em	ph:	
25:	em	pl:		25:	em	pl:	
26:	em	ro:		26:	em	ro:	
27:	em	sts:		27:	em	sts:	
28:	em	un:		28:	em	un:	
29:	em	vi:		29:	em	vi:	
30:	em	htg_hum:		30:	em	htg_hum:	
31:	em	htg_inv:		31:	em	htg_inv:	
32:	em	htg_other:		32:	em	htg_other:	
33:	em	htg_mus:		33:	em	htg_mus:	
34:	em	htg_pln:		34:	em	htg_pln:	
35:	em	htg_rtd:		35:	em	htg_rtd:	
36:	em	htg_nam:		36:	em	htg_nam:	
37:	em	htg_vrt:		37:	em	htg_vrt:	
38:	em	sy:		38:	em	sy:	
39:	em	htgo_hum:		39:	em	htgo_hum:	
40:	em	htgo_mus:		40:	em	htgo_mus:	
41:	em	htgo_other:		41:	em	htgo_other:	

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match		Length	DB	ID	Description
		Score	Match				
1	578.6	86.5	642	6	AX359705		AX359705 Sequence
2	564	84.3	648	9	BC009232		BC009232 Homo sapi
3	560.6	83.8	620	9	HSX318880		HSX318880 Homo sapi
4	549	82.1	580	6	AX078298		AX078298 Sequence
5	507.8	75.9	611	6	AX226501		AX226501 Sequence
6	287.2	42.9	493	9	HSX318881		AX318881 Homo sapi
C 7	245	36.6	20587	9	HSU18566A		Z88274 Human DNA
	245	36.6	62493	9	HSU193G15		AL117391 Human DNA
C 8	245	36.6	62493	9	HSU193G15		AL096838 Human DNA
C 10	245	36.6	145151	9	AL645949		AL645949 Human DNA
	216	32.3	2648	9	HSX318891		AX318891 Homo sapi
11	216	32.3	2648	9	HSX318891		AX318891 Homo sapi
12	216	32.3	94514	9	AL953874		AL953874 Human DNA
C 13	216	32.3	121324	9	EX293536		BL293536 Human DNA
14	215	32.1	320	6	AX198429		AX293536 Sequence
15	214	32.0	214	6	AX208978		AX198429 Sequence
16	204	30.5	399	6	AR272357		AX208978 Sequence
17	204	30.5	399	6	AR275338		AR272357 Sequence
18	204	30.5	399	6	AX062442		AX275338 Sequence
19	204	30.5	399	6	AX367359		AX062442 Sequence
20	204	30.5	457	6	AR273057		AX367359 Sequence
21	204	30.5	457	6	AR276638		AR273057 Sequence
22	204	30.5	457	6	AX063163		AR276638 Sequence
23	204	30.5	457	6	AX368080		AX063163 Sequence
24	204	30.5	463	6	AX370580		AX368080 Sequence
25	204	30.5	463	9	HSX290447		AX370580 Sequence
26	204	30.5	529	9	BC009538		AX290447 Homo sapi
27	204	30.5	626	9	AF251237		BC009538 Homo sapi
28	204	30.5	637	6	AX455517		AF251237 Homo sapi
29	204	30.5	740	9	HSX318878		AX455517 Sequence
30	204	30.5	767	6	BD109661		HSX318878 Homo sapi
31	204	30.5	781	6	AR273067		BD109661 EST and e
32	204	30.5	781	6	AR276648		AR273067 Sequence
33	204	30.5	781	6	AX063181		AR276648 Sequence
34	204	30.5	781	6	AX368098		AX063181 Sequence
35	182.8	27.3	528	9	AF055473		AX368098 Sequence
36	182.8	27.3	528	9	AF055473		AX275865 Sequence
37	181.2	27.1	530	9	HSU19143		AF055473 Homo sapi
38	181.2	27.1	538	6	AR028488		UI9143 Human GAGE-
39	181.2	27.1	538	6	BD132467		AR028488 Sequence
40	181.2	27.1	538	6	IS5851		BD132467 Isolated,
41	181.2	27.1	551	9	BC018052		IS5851 Sequence 14
42	180	26.9	479	6	AR272358		BC018052 Homo sapi
43	180	26.9	479	6	AR275939		AR272358 Sequence
44	180	26.9	479	6	AX062443		AR275939 Sequence
45	180	26.9	479	6	AX367360		AX062443 Sequence
46	179.6	26.8	1245	6	AX285022		AX367360 Sequence
47	178	26.6	461	6	AR273058		AX285022 Sequence
48	178	26.6	461	6	AR276639		AR273058 Sequence
49	178	26.6	461	6	AX063165		AR276639 Sequence
50	178	26.6	461	6	AX368082		AX063165 Sequence
51	178	26.6	481	9	HSX318979		AX368082 Sequence
52	169.8	25.4	524	9	HSU19146		AX318879 Homo sapi
53	169.8	25.4	527	9	HSU19147		AX318879 Homo sapi
54	169.8	25.4	528	6	AX334151		UI9146 Human GAGE-
55	169.8	25.4	528	9	HSU19145		UI9147 Human GAGE-
56	169.8	25.4	532	6	AR028491		UI9145 Human GAGE-
57	169.8	25.4	532	6	BD132470		AR028491 Sequence
58	169.8	25.4	532	6	IS5854		BD132470 Isolated,
59	169.8	25.4	539	6	AR028492		IS5854 Sequence 17
60	169.8	25.4	539	6	BD132471		AR028492 Sequence
61	169.8	25.4	539	6	IS5855		BD132471 Isolated,
62	169.8	25.4	540	6	AR028490		IS5855 Sequence 18
63	169.8	25.4	540	6	BD132469		AR028490 Sequence
64	169.8	25.4	540	6	IS5853		BD132469 Isolated,
65	169.8	25.4	637	9	BC024914		IS5853 Sequence 16
							BC024914 Homo sapi

cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
 http://www.systemsbio.org
 contact: amadan@systemsbiology.org
 Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
 Greene, Mark Kettman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 25 Row: m Column: 9
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, Similarity but not identity to protein.

FEATURES

```

source
1..648
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:16513 IMAGE:3960352"
/tissue_type="Placenta, choriocarcinoma"
/clone_lib="NIH_MGC_21"
/lab_host="DH10B-R"
/notes="Vector: pOT87"
203..538
/codon_start=1
/product="Similar to G antigen 8"
/protein_id="AAH09232.1"
/db_xref="GI:14328032"
/translators="MSWGRSTYRPRRSLOPPELIGAMLEPTDEPKKBPPTKSR
NPTFDQKREDQGAARIQVPLEADLQELCQTKTGCGEGTDVKGKILPKAEHFKMP
EAGEGKSOV"
BASE COUNT      207 a 134 c 162 g 145 t
ORIGIN
Query Match      84.3%; Score 564; DB 9; Length 648;
Best Local Similarity 95.8%; Pred. No. 6.2e-146;
Matches 598; Conservative 0; Mismatches 5; Indels 21; Gaps 1;

Qy 31 GTTCCTTGACACCTGCTCAGTGTGATGTTCACTGGGCATCTTCCCTCGACCCCTTT 90
Db 10 GTTCCTTGACACCTGCTCAGTGTGATGTTCACTGGGCATCTTCCCTCGACCCCTTT 69

Qy 91 GCCACGTGTGTACCGTGTGGGAGCTGTGAGTGTGAGGGGACAGTTCAGCCGCTCTGG 150
Db 70 GCCACGTGTGTACCGTGTGGGAGCTGTGAGTGTGAGGGGACAGTTCAGCCGCTCTGG 129

Qy 151 ACTCTTTCTCTCTACTGAGCGCAGCTTAGTTCGCGAGGCGAGTCTCCCGAGGAAT 210
Db 130 ACTCTTTCTCTCTACTGAGCGCAGCTTAGTTCGCGAGGCGAGTCTCCCGAGGAAT 189

Qy 211 GAATAGTGAATATGAGTTGGCGAGGAAGATCAACATATAGGCTTAGGCCAAGAAG 270
Db 190 GAATAGTGAATATGAGTTGGCGAGGAAGATCAACATATAGGCTTAGGCCAAGAAG 249

Qy 271 TTTCAGCTCTCTGAGCTGATTGGGCTATGCTTACTGGCTCCCTTTGTCCCGAGGAAC 330
Db 250 TTTCAGCTCTCTGAGCTGATTGGGCTATGCTTACTGGCTCCCTTTGTCCCGAGGAAC 288

Qy 331 CACTGATGAAGAGCTTAAGAGAGAAACCAACCCACTAAAGTCGGAATCTTACCTGA 390
Db 289 CACTGATGAAGAGCTTAAGAGAGAAACCAACCCACTAAAGTCGGAATCTTACCTGA 348

Qy 391 TCAGAGAGAGAGATGATGAGGTGCGAGTTCAGTTCCTAGTGCCTGACCTGGAGCCGA 450
Db 349 TCAGAGAGAGAGATGATGAGGTGCGAGTTCAGTTCCTAGTGCCTGACCTGGAGCCGA 408

Qy 451 TCTCAGAGAGTATCTCAGACAAAGACTGGGGATGGATGTGAAGGTGCTACTGATGCAA 510
Db 409 TCTCAGAGAGTATCTCAGACAAAGACTGGGGATGGATGTGAAGGTGCTACTGATGCAA 468

Qy 511 GGGGAAGATTCTACAAAGAGAGCACTTTAAATGCCAGAACAGGTGAAGGGAATC 570

```

```

Db 469 GGGGAAGATTCTACAAAGAGAGCACTTTAAATGCCAGAACAGGTGAAGGGAATC 528
Qy 571 ACAGGTTTAAAGAGATAAGCTGAAACAAACACAAACTGTTTTATATTAGATATTTTAC 630
Db 529 ACAGGTTTAAAGAGATAAGCTGAAACAAACACAAACTGTTTTATATTAGATATTTTAC 588
Qy 631 TTTAAAGAGCTTAAATAAATTTT 654
Db 589 TTTAAATAATCTTAAATAAGTTTT 612

RESULT 3
HSA318880
LOCUS      HSA318880      620 bp      mRNA      linear      PRI 14-JAN-2002
DEFINITION Homo sapiens mRNA for XAGE-2 protein.
ACCESSION  AJ318880.1 GI:18157209
VERSION     XAGE-2 gene.
KEYWORDS    Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1
AUTHORS    Zendman,A.J.W., van Kraats,A.A., Weidle,U.H., Ruiter,D.R. and Van
Muijen,G.N.P.
TITLE      Expression profile of members of the XAGE cancer/testis antigen
family
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 620)
AUTHORS    Zendman,A.J.W.
TITLE      Direct Submission
JOURNAL    Submitted (14-AUG-2001) Zendman A.J.W., Department of Pathology,
University Medical Center St Radboud, Geert Grooteplein Zuid 24, PO
Box 9101 6500 HB Nijmegen, NETHERLANDS

FEATURES
Source
1..620
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="Xp11.21-Xp11.22"
/tissue_type="Ewing sarcoma"
1..620
/gene="XAGE-2"
189..524
/gene="XAGE-2"
/codon_start=1
/product="XAGE-2 protein"
/protein_id="CAC83007.1"
/db_xref="GI:18157210"
/translators="MSWGRSTYRPRRSLOPPELIGAMLEPTDEPKKBPPTKSR
NPTFDQKREDQGAARIQVPLEADLQELCQTKTGCGEGTDVKGKILPKAEHFKMP
EAGEGKSOV"
BASE COUNT      191 a 131 c 156 g 142 t
ORIGIN
Query Match      83.8%; Score 560.6; DB 9; Length 620;
Best Local Similarity 96.0%; Pred. No. 5.5e-145;
Matches 594; Conservative 0; Mismatches 4; Indels 21; Gaps 1;

Qy 36 TTGGACACCTGCTCAGTGTGATGTTCACTGGGCATCTTCCCTCGACCCCTTTGCCCA 95
Db 1 TTGGACACCTGCTCAGTGTGATGTTCACTGGGCATCTTCCCTCGACCCCTTTGCCCA 60

Qy 96 CTTGTGTACCGCTGGGAGCTGTGAGTGTGAGGGGACAGTTCAGCCGCTGTGACTCT 155
Db 61 CGTGTGTACCGCTGGGAGCTGTGAGTGTGAGGGGACAGTTCAGCCGCTGTGACTCT 120

Qy 156 TTCTCTCTCTACTGAGACGCGAGCTTAGTTCGCGAGGCGAGTCTCCCGAGGAATCGAAAT 215
Db 121 TTCTCTCTCTACTGAGACGCGAGCTTAGTTCGCGAGGCGAGTCTCCCGAGGAATCGAAAT 180

Qy 216 AGTGAATATGATGTGGCGAGGAAGATCAACATATAGCCTAGGCCAAGAGAGTTTAC 275

```

```
Db 181 AGTGAATATGAGTGGCGAGGAGATCAATATAGAGCTAGCCAGGAAGAGTTTAC 240
Qy 276 AGCTCCTGAGCTGATTTGGGCTATGCTTACTGCTCCCTTTGTCCAGGAAGCCACTG 335
Db 241 AGCTCCTGAGCTGATTTGGGCTATGCTT-----GAAACCCACTG 279
Qy 336 ATGAAGAGCTTAAAGAGAGAAACACCCACTTAAAGTCGGAATCTTACACCTGATCAGA 395
Db 280 ATGAAGAGCTTAAAGAGAGAAACACCCACTTAAAGTCGGAATCTTACACCTGATCAGA 339
Qy 396 AGAGAGAGATGATCAGGCTGAGCTGAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 455
Db 340 AGAGAGAGATGATCAGGCTGAGCTGAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 399
Qy 456 AGGAGCTATGTGAGCAAGAGCTGGGATGAGATGTGAAGGTGGTACTGATGTCAAGGGGA 515
Db 400 AGGAGCTATGTGAGCAAGAGCTGGGATGAGATGTGAAGGTGGTACTGATGTCAAGGGGA 459
Qy 516 AGATTCTACAAAGCAGAGCTTTTAAATGCCAGAGCAGGTGAAGGGAATCAGAGG 575
Db 460 AGATTCTACAAAGCAGAGCTTTTAAATGCCAGAGCAGGTGAAGGGAATCAGAGG 519
Qy 576 TTTAAAGGAGATAGCTGAAACACACAACTGTTTTTATATAGATATTTTACTTTAA 635
Db 520 TTTAAAGGAGATAGCTGAAACACACAACTGTTTTTATATAGATATTTTACTTTAA 579
Qy 636 AGAGTCTTAATAATTTT 654
Db 580 AATATCTTAATAAGTTTT 598
```

```
RESULT 4
AX078298 AX078298 580 bp DNA linear PAT 22-FEB-2001
LOCUS Sequence 102 from Patent WO0107471.
DEFINITION AX078298
ACCESSION AX078298
VERSION AX078298.1 GI:13157989
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 Hillman,J.L., Lal,P., Tang,Y.T., Yue,H., Au-Young,J., Bandman,O.,
Azimzai,Y., Yang,J., Lu,D.A., Baughn,M.R., Patterson,C. and Shah,P.
TITLE Cell cycle and proliferation proteins
JOURNAL Patent: WO 0107471-A 102 01-FEB-2001;
Incyte Genomics, Inc. (US)
FEATURES Location/Qualifiers
source
1. 580
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/notice="Incyte ID No: 4764233CB1"
BASE COUNT 167 a 135 c 160 g 118 t
ORIGIN
```

```
Query Match 82.1%; Score 549; DB 6; Length 580;
Best Local Similarity 96.5%; Pred. No. 9.3e-142;
Matches 580; Conservative 0; Mismatches 0; Indels 21; Gaps 1;
Qy 1 CACAAGCAGGACCGACTTCAGTGTGCAATGTTCCCTTGACACCTCCCTCAGTGTGCA 60
Db 1 CACAAGCAGGACCGACTTCAGTGTGCAATGTTCCCTTGACACCTCCCTCAGTGTGCA 60
Qy 61 TTCACTGGGATCTTCCCTTCGACCCCTTTGCCACGATGGTGAACCGCTGGGAGCTGTGA 120
Db 61 TTCACTGGGATCTTCCCTTCGACCCCTTTGCCACGATGGTGAACCGCTGGGAGCTGTGA 120
Qy 121 GAGTGTGAGGGGACAGTTCAGCCGCTGACCTTTCTCTCTACTGAGCGCAGCTTA 180
Db 121 GAGTGTGAGGGGACAGTTCAGCCGCTGACCTTTCTCTCTACTGAGCGCAGCTTA 180
```

```
Qy 181 TAGGTCGCGAGGCCAGTCTCTCCAGGAATGAAATAGTGAATATAGTTCGCGAGGAG 240
Db 181 TAGGTCGCGAGGCCAGTCTCTCCAGGAATGAAATAGTGAATATAGTTCGCGAGGAG 240
Qy 241 ATCAACATATAGGCTTAGCCCAAGAGAGTTTACAGCTCTCTGAGCTGATTTGGGGCTAT 300
Db 241 ATCAACATATAGGCTTAGCCCAAGAGAGTTTACAGCTCTCTGAGCTGATTTGGGGCTAT 300
Qy 301 GCTTACTGGCTCCCTTTGTCTCCAGGAACCCACTGATGAAGAGCCCTTAAAGAGAGAAACC 360
Db 301 GCTT-----GACCCACTGATGAAGAGCCCTTAAAGAGAGAAACC 339
Qy 361 ACCCACTAAAGTCGAATCCTACACCTGATCAGAGAGAGAGATGATCAGGTCGACG 420
Db 340 ACCCACTAAAGTCGAATCCTACACCTGATCAGAGAGAGAGATGATCAGGTCGACG 399
Qy 421 TGAGATTCAGTGCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTTCAGACAAAGACTGG 480
Db 400 TGAGATTCAGTGCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTTCAGACAAAGACTGG 459
Qy 481 GATGATGTGAAGTGGTACTGATGTCAAGGGAGATTTACCAAAAGCAGACACTT 540
Db 460 GGATGATGTGAAGTGGTACTGATGTCAAGGGAGATTTACCAAAAGCAGACACTT 519
Qy 541 TAAATGCCAGAGCAGGTGAAGGGAATCAACAGGTTTAAAGGAAGATAGCTGAAACAA 600
Db 520 TAAATGCCAGAGCAGGTGAAGGGAATCAACAGGTTTAAAGGAAGATAGCTGAAACAA 579
Qy 601 C 601
Db 580 C 580
```

```
RESULT 5
AX226501 AX226501 611 bp DNA linear PAT 10-SEP-2001
LOCUS Sequence 7 from Patent WO0161009.
DEFINITION AX226501
ACCESSION AX226501
VERSION AX226501.1 GI:15555699
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 Malyankar,U.M., Tchernev,V.T., Padigaru,M., Taupier,R.J.,
Spytek,K.A., Majumder,K., Guo,X., Spaderna,S.K. and Boidog,F.L.
TITLE Polypeptides and nucleic acids encoding same
JOURNAL Patent: WO 0161009-A 7 23-AUG-2001;
Curagen Corporation (US)
FEATURES Location/Qualifiers
source
1. 611
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 196 a 127 c 152 g 135 t 1 others
ORIGIN
```

```
Query Match 75.9%; Score 507.8; DB 6; Length 611;
Best Local Similarity 94.6%; Pred. No. 2.8e-130;
Matches 566; Conservative 0; Mismatches 8; Indels 24; Gaps 3;
Qy 60 GTTCACTGGGATCTTCCCTTCGACCCCTTTGCCACGCTGGTGAACCCCT-GGGGAGCTGT 118
Db 9 GTCACTGGGATCTTCCCTTCGACCCCTTTGCCACGCTGGTGAACCCCTGGGAGCTGT 68
Qy 119 GAGAGTGTGAGGCGACGTTCCAGCCGCTGGACTTTCTCTCTCTACTGAGAGCGAGCC 178
Db 69 GAGAGTGTGAGGCGACGTTCCAGCCGCTGGACTTTCTCTCTCTACTGAGAGCGAGCC 128
Qy 179 TATAGTCCGAGCGCCAGTCTCCAGGAATAGTGAATATAGTTCGCGAGGAG 238
```

```
Db 129 TATAGTCCGAGCCAGTCTCCAGGAAGTGAATAGTGAATATGATGCTTGGCGAGGA 188
Qy 239 AGATCAACATATAGGCTAGCCAGGAAGAGTTTACAGCTCTCTGAGCTGATTTGGGGCT 298
Db 189 AGATCAACATATAGGCTAGCCAGGAAGAGTTTACAGCTCTCTGAGCTGATTTGGGGCT 248
Qy 299 ATGCTTACTGGCTCCCTTTGTCACAGAACCCACTGATGAAGAGCCCTAAAGAGAGAAA 358
Db 249 ATGCTT-----GAACCCACTGATGAGAGCCCTAAAGAGAGAAA 287
Qy 359 CCACCCCTAAAGTCGGAATCTTACACCTGA--TCAGAGAGAGAGAGATGATCAGGGTG 416
Db 288 CCACCCCTAAAGTCGGAATCTTACACCTGACTCAGAGAGAGAGAGATGATCAGGGTG 347
Qy 417 CAGCTGAGATTCAGTGCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGA 476
Db 348 CAGCTGAGATTCAGTGCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGA 407
Qy 477 CTGGGATGGATGGAAGTGGTACTGATGTCAGAGGGAAGATTTCTACCAAAAGCAGAGC 536
Db 408 CTGGGATGGATGGAAGTGGTACTGATGTCAGAGGGAAGATTTCTACCAAAAGCAGAGC 467
Qy 537 ACTTTAAATCCAGAGCAGTGAAGGAATCAGAGTTTAAAGGAAGTGAAGTGAAGTGA 596
Db 468 ACTTTAAATCCAGAGCAGTGAAGGAATCAGAGTTTAAAGGAAGTGAAGTGAAGTGA 527
Qy 597 ACAACACAACTGTTTTTATATTAGATATTTTACTTTAAAGAGCTTTTAAATATTTT 654
Db 528 ACAACACAACTGTTTTTATATTAGATATTTTACTTTAAATATCTTAAATAGTTT 585

RESULT 6
HSA318881 HSA318881 493 bp mRNA linear PRI 14-JAN-2002
LOCUS Homo sapiens mRNA for XAGE-3 protein.
DEFINITION
ACCESSION AJ318881
VERSION AJ318881.1 GI:18157211
KEYWORDS XAGE-3 gene.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE Zandman,A.J.W., van Kraats,A.A., Weidle,U.H., Ruiter,D.R. and Van
Muijen,G.N.P. Expression profile of members of the XAGE cancer/testis-antigen
family
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 493)
AUTHORS Zandman,A.J.W.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-2001) Zandman A.J.W., Department of Pathology,
University Medical Center St Radboud, Geert Grooteplein Zuid 24, PO
Box 9101 6500 HB Nijmegen, NETHERLANDS
FEATURES
Location/Qualifiers
source 1..493
/mol_type="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp11.21-Xp11.22"
/tissue_type="Ewing sarcoma"
1..493
/gene="XAGE-3"
61..396
/gene="XAGE-3"
/codon_start=1
/product="XAGE-3 protein"
/protein_id="CAC83008.1"
/db_xref="GI:18157212"
/translation="MWGRSTYRPRRSPVPPPELCPMLPEGDDEPQOEPPETSR
DPAGQREEDQGAETQVPLEADLQELSKTGGCGNPPDQGLKPSQFKNP
EGGDRQPV"
```

```
BASE COUNT 167 a 98 c 125 g 103 t
ORIGIN
Query Match 42.9%; Score 287.2; DB 9; Length 493;
Best Local Similarity 82.0%; Pred. No. 7.9e-69;
Matches 364; Conservative 0; Mismatches 58; Indels 22; Gaps 2;
Qy 213 ANTACTGAAATATGAGTTGGCGAGGAAGATCAACATATAGCCCTAGGCCACAGAGAACTT 272
Db 50 AAATGTGAAATATGAGTTGGCGAGGAAGATCAACATATAGCCCTAGGCCAGGAGAACTG 109
Qy 273 TACAGCCCTCTGAGCTGATTTGGGGCTATGCTTACTTGGCTCCCTTTGTGCCAGGAACCA 332
Db 110 TACCACCTCTGAGCTGATTTGGGGCTATGCT-----GGAGCCCG 148
Qy 333 CTGATGAAGAGCTTAAAGAGGAACACCACCACTTAAAGTCGGAATCCTACACTGATC 392
Db 149 GTGATGAGAGGCTCAGCAAGAGGAACCAACCACTGAAAGTCGGGATCCTGACCTGCTC 208
Qy 393 AGAAGAGAGGAAGATGATCAGGGTGCAGCTGAGATTCAGAGTGCCTGACCTGGAAGCCGATC 452
Db 209 AGGAGAGAGGAAGATGATCAGGGTGCAGCTGAGACTCAAGTGCCTGACCTGGAAGCTGATC 268
Qy 453 TCAGAGAGCTATGTCAGACAAAGACTGGGATCGATGTGAAGTGGTACTGATGTCAAG 512
Db 269 TCCAGAGAGCTGCTCAGTCAAGAGACTGGGGGTGAATGTGGAATGGTCTGATGACCCAGG 328
Qy 513 GGAAGATTTCAACAAAAGCAGAGCACTTTAAATGCCAGAGCAGGTGAAGGAAATCAC 572
Db 329 GGAAGATTTGCCAAAATCAGAACAAATTTAAATGCCAGAGGAGGTGACAGGCAACAC 388
Qy 573 AGGTTTAAAGGAAGATAGCTGAAACAACAC-AAAAGTCTTTTATATAGATATTTTACT 631
Db 389 AGGTTTAAATGAAGCAAGCTGAAACAACACAAACTGTTTTTATCTAAGATATTTGACT 448
Qy 632 TTAAGAGTCTTAAATAAATTTTGG 655
Db 449 TAAATATATCGAATTAACCTTTTG 472

RESULT 7
HSL18556A HSL18556A/20587 bp DNA linear PRI 23-NOV-1999
LOCUS Human DNA sequence from cosmid L129H7, Huntington's Disease Region,
DEFINITION chromosome 4p16.3 contains Pseudogene and CpG island.
ACCESSION 268274
VERSION 268274.1 GI:1130686
KEYWORDS 4p16.3; CpG island; pseudogene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 20587)
AUTHORS Buck,D.
TITLE Direct Submission
JOURNAL Submitted (12-DEC-1995) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, England. E-mail enquiries: humquery@sanger.ac.uk
COMMENT
IMPORTANT:
This sequence is not the entire insert of clone L18556. It may be
shorter because we only sequence overlapping sections once, or
longer because we arrange for a small overlap between neighbouring
submissions. In addition this sequence has been finished according
to sequence map criteria as follows. An attempt is made to resolve
all sequencing problems, such as compressions and repeats, but not
necessarily within known annotated human repeat sequence elements
(e.g. Alu). Where unambiguous sequence has not been obtained,
there is an annotation using the 'unsure' feature key. Where the
size of a problem region is known only approximately the sequence
contigs on either side will be submitted as separate entries, with
annotation to indicate the order and positions of neighbours and
what is known about the size and nature of the gap. Contigs have
entryname suffixes (A, B, ...) indicating the order of the contigs
with respect to the clone.
```

The sequence from clone L185B6 has been finished in more than one contig. This sequence (L185E6A) is separated from the following one (L185B6B) by a 900bp gap based on restriction data. 55bp VNTR made closure of gap impossible.

The true right end of clone L121P12 is at 10842 in this sequence. The true left end of clone L185B6 is at 1 in this sequence. L185B6 is from cosmid library LA04NC01 constructed at the Human Genome Center, Los Alamos National Laboratory, NM 87545 under the auspices of the U.S. Department of Energy. The library was constructed using flow-sorted human chromosome 4 from a Hamster-Human hybrid cell line (UV20HL21-27) containing human chromosomes 4, 8 and 21.

VECTOR: SCOS1

L185B6 is contained in a clone contig spanning 2mb which is described in Baxendale et al, Nature Genetics 4 (1993) 181- 186. See also Myers et al, Cytogenet Cell Genet. 66 (1994) 218-230.

FEATURES

```

Source
1..20587
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4p16.3"
/clone="LA04NC01-185B6"
/cell_line="UV20HL21-27"
/clone_lib="LA04NC01"
/complement(1370..1739)
/note="match: STS L00885"
2380..4066
/note="L1 element fragment"
4115..4400
/note="Alu repeat: matches 1..308 of consensus"
4427..4736
/note="L1 element fragment"
4778..5099
/note="L1 element fragment"
5203..5333
/note="MSTC element fragment"
5632..5721
/note="L1 element fragment"
5980..6714
/note="matches Trk-2h oncogene - 41aa of human RL7A_HUMAN M36072"
/codon_start=1
6747..6778
/note="15 copies of 2 mer 100 % conserved"
10906..11092
/partial
/note="Alu repeat: matches 1..191 of consensus"
11083..11214
/partial
/note="Alu repeat: matches 175..308 of consensus"
11352..11506
/partial
/note="Alu repeat: matches 308..151 of consensus"
13249..13799
/note="putative CpG island"
15946..16238
/note="Alu repeat: matches 1..308 of consensus"
17069..17143
/note="MLT1D element fragment"
17589..17689
/partial
/note="Alu repeat: matches 308..210 of consensus"
17695..17739
/partial
/note="Alu repeat: matches 170..114 of consensus"
18268..18369
/partial
/note="Alu repeat: matches 155..38 of consensus"
19668..19742
/note="MIR element fragment"

```

```

repeat_region 20313..20524
BASE COUNT 5439 a 5367 c 4601 g 5180 t
ORIGIN

```

```

Query Match 36.6%; Score 245; DB 9; Length 20587;
Best Local Similarity 70.9%; Pred. No. 6e-57;
Matches 426; Conservative 0; Mismatches 110; Indels 65; Gaps 5;
QY 58 ATGTTCACTGGGATCTCCCTTCGACCCCTTTGCCCACGTCGTCACCGCTGGGAGCTG 117
DB 18267 AGTTTCACTGGGATCTCTCTCTGCCCCCTTTGCCAGTGGTGAACACTGTGACCTG 18208
QY 118 TGAGAGTGTAGGGGACGCTTCAGCCGCTCTGAGACTTTTCTCTCTCTACTGAGCGCAGC 177
DB 18207 TGAGGGTGTAGGGGTTTCGTTCTCTGCGTCTTTTACCCACTTCTGAGAGCGAGC 18148
QY 178 CTATAGTCTCGCAGGCCAGTCTCTCCAGGAAGTGAATAGTGAATATGAGTTGGCGAGC 237
DB 18147 CGTT-----TGAAATATGAGTTGGCAAG 18124
QY 238 AGATCAACATATAGCCCTAGGCCAAGAAGAAGTTTACAGCCTCTCTGAGCTGATTCGGGC 297
DB 18123 AGATCAGCATGTAGGCTTTAGACCAAGACGCTATGTGCAGCCTCTCTGAGCTAACCTGGGC 18064
QY 298 TATGCTTACTGGCTCCCTTTGTCCAGGAACCCACTGATGAAGAGCCCTAAAGAGAGAA 357
DB 18063 TGTGCTT-----GAGCCAGTGTAGCAGCCTCAGCAAGAGGA 18025
QY 358 ACCACCCACTAAAGTCGGAATCTTACACTGATCAGAGAGAGAAGATGATCAGGTCG 417
DB 18024 ACCACCACTGAAAGTCGGGGTCTT--ACCTGGCCAGGAGAGAGGAAGATCAGGTCG 17967
QY 418 AGCTGAGATTTCAAGTGCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTTCAGCAAGAC 477
DB 17966 TGTGAGATTTCTGTGCTTGACCAAGAGCTGATCTCCGGAGCTCTCTCAAGA----- 17912
QY 478 TGGGATGAGTGTAAAGTGGTACTGATGTCAAGGGGAAGATTCTTACCAAGAGCAGACA 537
DB 17911 CTGGGATGAATGTGGAGATGGTCTCTGATGTC-CGGGGAATATTCTCCGGAATCAGACA 17853
QY 538 CTTTAAATCCAGAGAGCGGTGAAGGAAATCAGAGGTTTAAAGGAAGATAGCTGAAA 597
DB 17852 ATTAAATGCCAGAGAGGTTGAAGGCAACACAGGTTCCAGAGAGCAGAGCTGAAA 17793
QY 598 CAACACAACTGTTTTTATATTAGATATTTTACTTTTAAAGAGTCTTAAATAATTTTGGC 657
DB 17792 CAATGCAAACTGGTTTTTATATTAGTACGTGACTTAAATATCTCAATACAGTTTCTCC 17733
QY 658 A 658
DB 17732 A 17732

```

```

RESULT 8
LOCUS HS193G15 62493 bp DNA linear PRI 05-MAR-2003
DEFINITION Human DNA sequence from clone RP1-193G15 on chromosome 4, complete
sequence.
ACCESSION AL117391
VERSION AL117391.1 GI:5911814
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 62493)
McMurray,A.
Direct Submission
Submitted (05-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
hmuqy@sanger.ac.uk Clone requests: clonersrequest@sanger.ac.uk
----- Genome Center
COMMENT

```

Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep/RP1-193G15 is from the library RPCI-1 constructed by the group of Pieter de Jong. For further details see

<http://www.choiri.org/bacpac/home.htm>
VECTOR: pCYPAC2.

FEATURES

source
1. 62493
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/clone="RP1-193G15"
/clone_lib="RPCI-1"
BASE COUNT 17336 a 16390 c 13580 g 15187 t
ORIGIN
Query Match 36.6%; Score 245; DB 9; Length 62493;
Best Local Similarity 70.9%; Pred. No. 6.5e-57;
Matches 426; Conservative 0; Mismatches 110; Indels 65; Gaps 5;
QY 58 ATGTTCACTGGGCACTTCCTTCGACCCCTTTGCCACGGTGACCGCTGGGGAGCTG 117
Db 54358 AGGTTCACTGGGTCTCTCTCTGCGCCCTTTGCCACGGTGTAACACTGTGGACTG 54299
QY 118 TGAGAGTGTAGGGGCAAGTTCACGCCCTCTGGACTCTTTCTCTCTACTGAGACGAGC 177
Db 54298 TGAGGGTGTAGGGGTGCGTTCCTCGCTCTGGACTCTTTTACCCCACTACTGAGAAGCAGC 54239
QY 178 CTATAGGTCCGACGCCAGTCTCCACGGAACCTGAATAGTGAATATGAGTTGGCGAGG 237
Db 54238 CGTT-----TGAATATGAGTTGGCAAGG 54215
QY 238 AAGATCAATATAGGCTAGGCCAAGAAAGTTTACAGCTCTCGAGCTGATGGGCGC 297
Db 54214 AAGATCAGCATGTAGGCTTAGACCAAGACGGTATGTGACGCTCTCGAGCTTAAGTGGCC 54155
QY 298 TATGCTTACTGGCTCCCTTTGTCCAGGACCCACTGATGAAGAGCTTAAGAGAGAA 357
Db 54154 TGTGTT-----GAGCCAGTATGAGCAGCTCTAGCAAGAGGA 54116
QY 358 ACCACCCCTAAAAGTCGGAATCTCTACACCTGATCAGAAGAGAGAAGATGATCAGGGTGC 417
Db 54115 ACCACCACTGAAGTCGGGTCTT--ACCTGGCCAGGAGAGAGAGATCAGCTGC 54058
QY 418 AGCTGAGATCAAGTCCCTGACCTGAGAGCGATCTCCAGAGCTATGTGACAAAGAC 477
Db 54057 TGCTGAGATCTTGTGCTGTGACCAAGAGCTGATCTCGGGAGCTGTCTCAAGA----- 54003
QY 478 TGGGATGATCTGAAGTGGTGTACTGTGTCAAGGGGAAGATTCTTACCAAGCAGAGCA 537
|||||

Db 54002 CTGGGATGAATGTGGAGATGGTCTGTGATGTC-CGGGGATATTCTCCGAATCAGACA 53944
QY 538 CTTTAAATGCCAGAGCGAGTGAAGGGAATCAACAGGTTTAAAGGAAGATAGCTGAA 597
|||||
Db 53943 ATTAAATGCCAGAGGAGGTGAAGGGCAACACAGGTTCCAGGAGACACAGCTGAA 53884
|||||
QY 598 CAACACAAACTGTTTTATATTAGATATTTTAAAGAGTCTTAATAAATTTTGGC 657
Db 53883 CAATGAAACTGTTTTATATTAGATACGTGACTTAAATAATCTCAATACAGTTTCTCC 53824
QY 658 A 658
Db 53823 A 53823

RESULT 9

HSJ193G15/c

LOCUS

DEFINITION

Human DNA sequence from clone 193G15 on chromosome 4

5', part of the RGS12 gene for regulator of G-protein signalling 12, an RPL7A (60S Ribosomal Protein 7A (SURF3)) pseudogene, ESTs, an STS, GSSs and a putative CpG island, complete sequence.

ACCESSION

AL096838.1 GI:5459253

VERSION

HTG; 60S Ribosomal Protein 7A; CpG Island; G-protein; RGS12; RPL7A; SURF3.

KEYWORDS

Surf3.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 62493)

McMurray, A.

Direct Submission

Submitted (10-JUL-1999) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

requests: clonerequest@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

The following abbreviations are used to associate primary accession

numbers given in the feature table with their source databases:

Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL

IMPORTANT: This sequence is not the entire insert of clone 193G15.

It may be shorter because we only sequence overlapping sections

once, or longer because we arrange for a small overlap between

neighbouring submissions.

The true left end of clone 1E6 is at 6371 in this sequence. The

true right end of clone 361H4 is at 104 in this sequence. This

sequence has been finished according to sequence map criteria as

follows. An attempt is made to resolve all sequencing problems,

such as compressions and repeats, but not necessarily within known

annotated human repeat sequence elements (e.g. Alu). Where the

sequence is ambiguous, there is an annotation using the 'unsure'

feature key.

This sequence was generated from part of bacterial clone contigs of

human chromosome 4, constructed by the Sanger Centre Chromosome 4

Mapping Group.

193G15 is from the library RPCI1 constructed at the Roswell Park

Cancer Institute by the group of Pieter de Jong. For further

details see <http://bacpac.med.buffalo.edu/VECTOR:pcYPAC2>.

Location/Qualifiers

1. 62493

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="4"

/clone="RP1-193G15"

/clone_lib="RPCI-1"

BASE COUNT 17336 a 16390 c 13580 g 15187 t

ORIGIN


```

Db      101976 ATTTAAATTGCCAGAGGAGTGGAAGCGCACACAGGTTCCAGGAGAACAGCTGAAA 101917
QY      598 CACACAAACGTTTTATATTAGATATTTTACCTTTAAAGAGTCTTTAATAATTTTGGC 657
Db      101916 CAATGCAACATGGTTTATTATTAGATACGTTAAATATCTCAATACAGTTTCTCC 101857
QY      658 A 658
Db      101856 A 101856

RESULT 11
HSA318891
LOCUS   HSA318891      2648 bp      DNA      linear      PRI 15-JAN-2002
DEFINITION Homo sapiens partial XAGE-2 gene for XAGE-2 5' part, exons 1-2 (and joined CDS).
ACCESSION AJ318891
VERSION   AJ318891.1 GI:18182816
KEYWORDS XAGE-2 gene.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1. Zandman, A.J.W., van Kraats, A.A., Weidle, U.H., Ruiter, D.J. and Van Muijen, G.N.P. Expression profile and alignment of the XAGE family of cancer/testis associated genes Unpublished
AUTHORS   Zandman, A.J.W.
TITLE     Direct Submission
JOURNAL   Submitted (15-AUG-2001) Zandman A.J.W., Department of Pathology, University Medical Center, Geert Grooteplein Zuid 24, P.O. Box 9101 6500 HB Nijmegen, NETHERLANDS
COMMENT   Related sequence AJ318892.
FEATURES  Location/Qualifiers
           1. 2648
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"
               /chromosome="X"
               /map="Xp11.21-Xp11.22"
               /join="(1441..2601,AJ318892.1:407..5517)
gene      join(1441..1720,2513..2601,AJ318892.1:407..512,
mRNA      AJ318892.1:2384..2509,AJ318892.1:5409..5517)
           /gene="XAGE-2"
           /label="XAGE-2_mRNA"
           1441..1720
           /gene="XAGE-2"
           /number=1
           1721..2512
           /gene="XAGE-2"
           /number=1
           2513..2601
           /gene="XAGE-2"
           /number=2
           join(2521..2601,AJ318892.1:407..512,AJ318892.1:2384..2509,
           AJ318892.1:5409..5431)
           /gene="XAGE-2"
           /codon_start=1
           /label="XAGE-2_CDS"
           /product="XAGE-2 protein"
           /protein_id="CAC88125.1"
           /db_xref="GI:18182817"
           /translation="NSWRGRSTYRPRRSLQPELIGAMLEPTDEEPKKEKPTKSR
           NPTDKRRDQGAATIGVPLEADLQLCQTKTGDCGCGTDVKGKILPAEHHFKNP
           EAGEGRSQV"
BASE COUNT 731 a 557 c 717 g 643 t
ORIGIN
Query Match 32.3%; Score 216; DB 9; Length 2648;

```

```

Best Local Similarity 93.8%; Pred. No. 6.1e-49;
Matches 225; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY      1 CACAACGAGGACGACGACTTCAAGTGTGCATGTTCTTGGACACCTGCTCAGTGTGCATG 60
Db      1506 CACAACGAGGACGACGACTTCAAGTGTGCATGTTCTTGGACACCTGCTCAGTGTGCATG 1565
QY      61 TTCACTGGGCAATCTTCCCTTTCGACCCCTTTGCCACGCTGGTGACCGCTGGGGAGGCTGTGA 120
Db      1566 TTCACTGGGCAATCTTCCCTTTCGACCCCTTTGCCACGCTGGTGACCGCTGGGGAGGCTGTGA 1625
QY      121 GAGTGTGAGGGGACGCTTCCAGCCGCTCTGGACTCTTTCTCTCTACTGAGACGACGCTTA 180
Db      1626 GAGTGTGAGGGGACGCTTCCAGCCGCTCTGGACTCTTTCTCTCTACTGAGACGACGCTTA 1685
QY      181 TAGGTCGCGAGCCAGCTCTCCAGCAAGTGAATAGTGAATATAGTGGCGAGGAGG 240
Db      1686 TAGGTCGCGAGCCAGCTCTCCAGCAAGTGAATAGTGAATAGTGGCGAGGAGG 1745

RESULT 12
AL953874
LOCUS   AL953874      94514 bp      DNA      linear      PRI 01-NOV-2002
DEFINITION Human DNA sequence from clone CTD-2267G17 on chromosome X, complete sequence.
ACCESSION AL953874
VERSION   AL953874.2 GI:24527895
KEYWORDS HTG.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1. (bases 1 to 94514)
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
           Wray, P.
           Direct Submission
           Submitted (01-NOV-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
           humquery@sanger.ac.uk Clone requests: clonesrequest@sanger.ac.uk
           On Nov 3, 2002 this sequence version replaced gi:24410496.
COMMENT   ----- Genome Center
           Center: Wellcome Trust Sanger Institute
           Center code: SC
           Web site: http://www.sanger.ac.uk
           Contact: humquery@sanger.ac.uk
           -----
           During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
           This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at
           http://www.sanger.ac.uk/Projects/C.elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at
           http://www.sanger.ac.uk/HGP/ChrX
           CTD-2267G17 is from the Caltech genomic sperm BAC library D. VECTOR: pBelobAC11.
FEATURES  Location/Qualifiers
           1..94514
               /organism="Homo sapiens"
               /mol_type="genomic DNA"

```


Db 61 GGGATGGATGTAAGGTGGTACTGATGTCAGGGGAAGATTCTACCAAAAGCAGAGCACT 120
Qy 540 TTAATAATGCCAGACAGGTCAGGGGAATCACAGGTTTAAAGAGATAGCTGGAACA 599
Db 121 TTAATAATGCCAGACAGGTCAGGGGAATCACAGGTTTAAAGAGATAGCTGGAACA 180
Qy 600 ACACAACTGTTTATATATAGATATTTACTTTA 634
Db 181 ACACAACTGTTTATATATAGATATTTACTTTA 215
RESULT 15
LOCUS AX208978 214 bp DNA linear PAT 31-AUG-2001
DEFINITION Sequence 818 from Patent WO0157207.
ACCESSION AX208978
VERSION AX208978.1 GI:15423401
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Algate, P.A. and Mannion, J.
TITLE Compositions and methods for the therapy and diagnosis of ovarian cancer
JOURNAL Patent: WO 0157207-A 818 09-AUG-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1..214
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 74 a 34 c 55 g 51 t
ORIGIN
Query Match 32.0%; Score 214; DB 6; Length 214;
Best Local Similarity 100.0%; Pred. No. 1.8e-48;
Matches 214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 420 CTGAGATTCAAGTGGTACCTGGAGCGGATCTCCAGGAGCTATGTCAGACAAAGACTG 479
Db 1 CTGAGATTCAAGTGGTACCTGGAGCGGATCTCCAGGAGCTATGTCAGACAAAGACTG 60
Qy 480 GGGATGGATGTAAGGTGGTACTGATGTCAGGGGAAGATTCTACCAAAAGCAGAGCACT 539
Db 61 GGGATGGATGTAAGGTGGTACTGATGTCAGGGGAAGATTCTACCAAAAGCAGAGCACT 120
Qy 540 TTAATAATGCCAGACAGGTCAGGGGAATCACAGGTTTAAAGAGATAGCTGGAACA 599
Db 121 TTAATAATGCCAGACAGGTCAGGGGAATCACAGGTTTAAAGAGATAGCTGGAACA 180
Qy 600 ACACAACTGTTTATATATAGATATTTACTTTT 633
Db 181 ACACAACTGTTTATATATAGATATTTACTTTT 214
RESULT 16
LOCUS AR272357 399 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 69 from patent US 6504010.
ACCESSION AR272357
VERSION AR272357.1 GI:29704242
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 399)
AUTHORS Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S.,
Carter, D., Retter, M.W., Mannion, J., Fan, L., and Wang, A.
TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: US 6504010-A 69 07-JAN-2003;
FEATURES Location/Qualifiers
source 1..399
/organism="unknown"
BASE COUNT 150 a 87 c 94 g 67 t 1 others
ORIGIN
Query Match 30.5%; Score 204; DB 6; Length 399;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
Qy 319 GTCCAGGAAACCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCCTAAAGTCGGAA 378
Db 40 GTCCAGGAGCCCGATGATGGAGAGCCCAAAAGAGAGAACCCAGCAGCTGAAAGTCGGGA 99
Qy 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAGTGCCTGA 438
Db 100 TCCTACACCTGGGCAGCAGACAGAGAGATCAGGTAGATCCCGAGTCCCGAGTCGGGA 159
Qy 439 CCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGGATGATGTAAGGT-G 497
Db 160 CATGGAGAGTGTCTGCAAGAGCTGCATCAGTCAAAACCCGGGATAAATCTGGATTGG 219
Qy 498 GTACTGATGTCAAGGGGAAGATTCTACAAAAGCAGAGCATTAAATGCCAGAGCAG 557
Db 220 GTTCCGCGCTCAAGGTGAAGATAATACCTTAAAGAGAGACACTGTAAATGCCAGAGCAG 279
Qy 558 GTCAAGGGAAATCACAGGTTTAAAGGAAGATAGCTGAAACCAACACAAACTGTTTTATA 617
Db 280 GTGAGAGCAACCAAGTTTAAATGAAGACAGCTGAACCAACCGCAAGCTGTTTTATA 339
Qy 618 TTAGATATTTTACTTTTAAAGAGTCTTTAAATTTT 653
Db 340 TTAGATATTTTACTTTTAAAGTATCTCAATAAAGTTTT 375
RESULT 17
LOCUS AR275938 399 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 69 from patent US 6509448.
ACCESSION AR275938
VERSION AR275938.1 GI:29709583
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 399)
AUTHORS Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S.,
Carter, D., Retter, M.W., Mannion, J., Fan, L., and Wang, A.
TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: US 6509448-A 69 21-JAN-2003;
FEATURES Location/Qualifiers
source 1..399
/organism="unknown"
BASE COUNT 150 a 87 c 94 g 67 t 1 others
ORIGIN
Query Match 30.5%; Score 204; DB 6; Length 399;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
Qy 319 GTCCAGGAAACCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCCTAAAGTCGGAA 378
Db 40 GTCCAGGAGCCCGATGATGGAGAGCCCAAAAGAGAGAACCCAGCAGCTGAAAGTCGGGA 99
Qy 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAGTGCCTGA 438
Db 100 TCCTACACCTGGGCAGCAGACAGAGAGATCAGGTAGATCCCGAGTCCCGAGTCGGGA 159
Qy 439 CCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGGATGATGTAAGGT-G 497
Db 160 CATGGAGAGTGTCTGCAAGAGCTGCATCAGTCAAAACCCGGGATAAATCTGGATTGG 219

```
QY 498 GTACTGATGTCAGGGGAGATTCTACCAAGACGAGACACTTTAAATGCCAGAGCAG 557
Db 220 GTTCGGGGTCAAGTGAAGTAATACCTTAAGAGGAACTGTAAATGCCAGAGCAG 279
QY 558 GTGAAGGGAATCAGAGGTTTAAAGGAGATAGCTGAAACAAACAACTGTTTTTATA 617
Db 280 GTGAAGAGCAACCAAGTTTAAATGAAGACAGCTGAAACACCAAGCTGGTTTTATA 339
QY 618 TTAGATATTTTACTTTAAAGAGCTCTTAATAATTTT 653
Db 340 TTAGATATTTGACTTAAACTATCTCAATAAAGTTTT 375

RESULT 18
AX062442 399 bp DNA linear PAT 24-JAN-2001
LOCUS Sequence 69 from Patent WO0100828.
ACCESSION AX062442
VERSION AX062442.1 GI:12540317
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S.,
Carter, D., Retter, M.W. and Mannion, J.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0100828-A 69 04-JAN-2001;
CORIXA CORPORATION (US)
FEATURES
source
1. .399
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 150 a 87 c 94 g 67 t 1 others
ORIGIN
Query Match 30.5%; Score 204; DB 6; Length 399;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGGACCCACTGATGAAGAGCTTAAGAGAGAACACCCACCTAAAGTCCGGA 378
Db 40 GTCCAGGAGGCCAGTAAATGGAGAGCCCCAAAGAGAACCCAGCAGCTGAAGTCCGGA 99
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGA 438
Db 100 TCCTACACCTGGCAGCAGACAGAGAGATCAGATACAGCTGAGATCCAGTGCAGCA 159
QY 439 CCTGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGCTGGGATGGATGAGT-G 497
Db 160 CATGGAAGTGTATCTGCAAGAGCTGCATCAGTCAACACCGGGGATAAATCTGATTGG 219
QY 498 GTACTGATGTCAGGGGAGATTCTACCAAGACGAGACACTTTAAATGCCAGAGCAG 557
Db 220 GTTCGGGGTCAAGTGAAGTAATACCTTAAGAGGAACTGTAAATGCCAGAGCAG 279
QY 558 GTGAAGGGAATCAGAGTTTAAAGGAGATAGCTGAAACAAACAACTGTTTTTATA 617
Db 280 GTGAAGAGCAACCAAGTTTAAATGAAGACAGCTGAAACACCAAGCTGGTTTTATA 339
QY 618 TTAGATATTTTACTTTAAAGAGCTCTTAATAATTTT 653
Db 340 TTAGATATTTGACTTAAACTATCTCAATAAAGTTTT 375

RESULT 19
AX367359 399 bp DNA linear PAT 16-FEB-2002
LOCUS Sequence 69 from Patent WO0204514.
DEFINITION AX367359
ACCESSION AX367359
```

```
VERSION AX367359.1 GI:18855463
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM
REFERENCE
AUTHORS Wang, T., Watanabe, Y., Henderson, R.A., Johnson, J.C., Retter, M.W.,
Mannerakis, M., Carter, D., Fanger, G.R., Vedvick, T.S., Bangur, C.S.,
McNabb, A., Fanger, N., Switzer, A., McNeill, P.D. and Clapper, J.D.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0204514-A 69 17-JAN-2002;
CORIXA CORPORATION (US)
FEATURES
source
1. .399
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 150 a 87 c 94 g 67 t 1 others
ORIGIN
Query Match 30.5%; Score 204; DB 6; Length 399;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGGACCCACTGATGAAGAGCTTAAGAGAGAACACCCACCTAAAGTCCGGA 378
Db 40 GTCCAGGAGGCCAGTAAATGGAGAGCCCCAAAGAGAACCCAGCAGCTGAAGTCCGGA 99
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGA 438
Db 100 TCCTACACCTGGCAGCAGACAGAGAGATCAGATACAGCTGAGATCCAGTGCAGCA 159
QY 439 CCTGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGCTGGGATGGATGAGT-G 497
Db 160 CATGGAAGTGTATCTGCAAGAGCTGCATCAGTCAACACCGGGGATAAATCTGATTGG 219
QY 498 GTACTGATGTCAGGGGAGATTCTACCAAGACGAGACACTTTAAATGCCAGAGCAG 557
Db 220 GTTCGGGGTCAAGTGAAGTAATACCTTAAGAGGAACTGTAAATGCCAGAGCAG 279
QY 558 GTGAAGGGAATCAGAGTTTAAAGGAGATAGCTGAAACAAACAACTGTTTTTATA 617
Db 280 GTGAAGAGCAACCAAGTTTAAATGAAGACAGCTGAAACACCAAGCTGGTTTTATA 339
QY 618 TTAGATATTTTACTTTAAAGAGCTCTTAATAATTTT 653
Db 340 TTAGATATTTGACTTAAACTATCTCAATAAAGTTTT 375

RESULT 20
AX367359 457 bp DNA linear PAT 10-APR-2003
LOCUS Sequence 790 from patent US 6504010.
DEFINITION
ACCESSION AR273057
VERSION AR273057.1 GI:29704942
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 457)
AUTHORS Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S.,
Carter, D., Retter, M.W., Mannion, J. and Fan, L.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: US 6504010-A 790 07-JAN-2003;
FEATURES
source
1. .457
/organism="unknown"
BASE COUNT 146 a 105 c 121 g 85 t
ORIGIN
```


1
 Wang, T., Watanabe, Y., Henderson, R. A., Johnson, J. C., Retter, M. W.,
 Marnerakis, M., Carter, D., Fanger, G. R., Vedvick, T. S., Bangur, C. S.,
 McNabb, A., Fanger, N., Switzer, A., McNeill, P. D. and Clapper, J. D.
 Compositions and methods for the therapy and diagnosis of lung
 cancer
 Patent: WO 0204514-A 790 17-JAN-2002;
 CORIXA CORPORATION (US)
 Location/Qualifiers
 1. 457
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 146 a 105 c 121 g 85 t
 BASE COUNT
 ORIGIN

Query Match 30.5%; Score 204; DB 6; Length 457;
 Best Local Similarity 77.4%; Pred. No. 1.2e-45;
 Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

Qy 319 GTCCAGGAACCCACTGATGAAGAGCCCTCAAGAGAGAGAAACCCCACTAAAAGTCGGAA 378
 Db 120 GTCCAGGAGCCCAAGTAATGGAGAGCCCAAAAGAAAGAACCCAGCAGCTGAAAGTCGGGA 179
 Qy 379 TCCTACACCTCATCAGAGAGAGAGAGATGATCAGGGTCAGCTGAGATTCAGATGCGCTGA 438
 Db 180 TCCTACACCTGGCGCAGCAGACAGAAAGATCAGATACAGCTGAGATCCAGTGC CGGA 239
 Qy 439 CCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGATCGATGTGAAGGT-G 497
 Db 240 CATGGAAGGTGATCTGCAAGAGCTGCGATCAGTCAAACACCGGGGATAAATCTGGATTGG 299
 Qy 498 GTACTGATGTCAGAGGGAAGATTCTACCAAAGCAGAGCACTTTAAAAATGCCAGAAGCAG 557
 Db 300 GTTCGGCGCTCAAGGTGAAGATATACCTAAAGAGGACACTGTAAATGCCAGAAGCAG 359
 Qy 558 GTGAAGGGAATCAAGGTTTAAAGGAGATAGCTGAACACACACAACTGTTTTTATA 617
 Db 360 GTGAAGAGCAACCAAGATTTAAATGAAGACAAGCTGAACCAAGCTGAAACCAAGCTGTTTTTATA 419
 Qy 618 TTGATATTTTACTTTAAAGAGTCTTAAATAATTTT 653
 Db 420 TTGATATTGACTTAACTATCTCAATAAAGTTT 455

RESULT 24
 AX370580
 LOCUS
 DEFINITION Sequence 1 from Patent EP1179589.
 ACCESSION AX370580
 VERSION AX370580.1 GI:19168789
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1

REFERENCE
 AUTHORS
 TITLE Mmx-1, a member of the family of human cancer/testis antigens, a
 protein encoded thereby and a process for determining whether a
 tumor sample has metastatic potential
 Patent: EP 1179589-A 1 13-FEB-2002;
 F. HOFFMANN-LA ROCHE AG (CH)
 Location/Qualifiers
 1. 463
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 126. 371
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD26691.1"
 /db_xref="GI:19168790"

JOURNAL
 FEATURES
 source

CDS

```
/protein_id="CAC38107.1"
/db_xref="GI:1392499"
/translation="MESPKKKNQQLKVGILHLSRQKIRIQLRSQCATWKVICKSCI
SQTGINDLGSVGVKLIIPKEEHCKMPEAGEBPQV"
BASE COUNT 155 a 106 c 120 g 82 t
ORIGIN
Query Match 30.5%; Score 204; DB 9; Length 463;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGGAACCCACTGATGAAGAGCCTTAAGAGAGAGAAACCCACCTAAAGTCGGAA 378
Db 108 GTCCAGGAGCCAGTAATGAGAGCCCAAAAGAGAGAAACCCAGCAGCTGAAGTCGGGA 167
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCAGCTGAGATTCAGTGCCTGA 438
Db 168 TCCTACACCTGGCAGCAGCAGAGAGATCAGATCAGCTGAGATCCAGTGCCTGA 227
QY 439 CTTGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGATGGATGAGGT-G 497
Db 228 CATGGAAGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGG 287
QY 498 GTACTGATGTCNAGGGAGGATTCTACCAAGCAGAGCAGCTTTAAATGSCCAGAGCAG 557
Db 288 GTTCGGCGCTCAAGGTGAAGATAATACCTAAAGAGGAACTGTAAATGCCAGAGCAG 347
QY 558 GTGAAGGGAATCAGAGCTTTAAAGAGAGATAAGCTGAACCAACACACAACTGTTTTATA 617
Db 348 GTGAAGAGCAACCAACAGTTTAAATGAAGACAAGCTGAACCAACCGCAGCTGTTTTATA 407
QY 618 TTAGATATTTTACTTTAAAGAGCTCTTAATAATTTT 653
Db 408 TTAGATATTTGACTTTAAACTATCTCAATAAAGTTT 443
RESULT 26
BC009538 529 bp mRNA linear PRI 22-OCT-2001
LOCUS Homo sapiens, clone MGC:10063 IMAGE:3893227, mRNA, complete cds.
DEFINITION BC009538
ACCESSION BC009538
VERSION BC009538.1 GI:16306934
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Strausberg, R.
Direct Submission
Submitted (29-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTB/Drp/Gazdar
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 14 Row: 9 Column: 8.
Location/Qualifiers
1. 529
/organism="Homo sapiens"
FEATURES
source
```

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:10063 IMAGE:3893227"
/tissue_type="lung, carcinoma, large cell
undifferentiated."
/clone_lib="NIH MGC_69"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
184..429
/codon_start=1
/product="Unknown (protein for MGC:10063)"
/db_protein_id="AAH09538.1"
/db_xref="GI:16306935"
/translation="MESPKKKNQQLKVGILHLSRQKIRIQLRSQCATWKVICKSCI
SQTGINDLGSVGVKLIIPKEEHCKMPEAGEBPQV"
BASE COUNT 184 a 116 c 146 g 83 t
ORIGIN
Query Match 30.5%; Score 204; DB 9; Length 529;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGGAACCCACTGATGAAGAGCCTTAAGAGAGAGAAACCCACCTAAAGTCGGAA 378
Db 167 GTCCAGGAGCCAGTAATGAGAGCCCAAAAGAGAGAAACCCAGCAGCTGAAGTCGGGA 226
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCAGCTGAGATTCAGTGCCTGA 438
Db 227 TCCTACACCTGGCAGCAGCAGAGAGATCAGATCAGCTGAGATCCAGTGCCTGA 286
QY 439 CTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGATGGATGGAAGGT-G 497
Db 287 CATGGAAGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGG 346
QY 498 GTACTGATGTCNAGGGAGGATTTCTACCAAGCAGAGCAGCTTTAAATGCCAGAGCAG 557
Db 347 GTTCGGCGCTCAAGGTGAAGATAATACCTAAAGAGGAACTGTAAATGCCAGAGCAG 406
QY 558 GTGAAGGGAATCAGAGCTTTAAAGAGAGATAAGCTGAACCAACACCAACTGTTTTATA 617
Db 407 GTGAAGGCAACCAACAGTTTAAATGAAGACAAGCTGAACCAACCGCAGCTGTTTTATA 466
QY 618 TTAGATATTTTACTTTAAAGAGCTCTTAATAATTTT 653
Db 467 TTAGATATTTGACTTTAAACTATCTCAATAAAGTTT 502
RESULT 27
AF251237 626 bp mRNA linear PRI 23-AUG-2000
LOCUS Homo sapiens XAGE-1 mRNA, complete cds.
DEFINITION AF251237
ACCESSION AF251237
VERSION AF251237.1 GI:9885324
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 626)
XAGE-1, A New Gene That is Frequently Expressed in Ewing's Sarcoma
Unpublished
2 (bases 1 to 626)
REFERENCE
AUTHORS Liu, X.F., Helman, L.J., Yeung, C., Lee, B. and Pastan, I.
TITLE Direct Submission
AUTHORS Liu, X.F., Helman, L.J., Yeung, C., Lee, B. and Pastan, I.
TITLE Submitted (31-MAR-2000) Lab of Mol. Biol., NCI, NIH, 37 Convent Dr.
JOURNAL MSC 4255, Bethesda, MD 20892, USA
FEATURES
source
1. 626
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X"
```


/gene="XAGE-1c"
/codon_start=1
/product="XAGE-1c protein"
/protein_id="CAC82986.1"
/db_xref="GI:18157206"
/translation="MRCHAGPSCLVTAITREGGPRSGGAQAALGCGWGPSPRSTW
NPDREFTWPTQPGGHERHTOTQNHPTASPRSPMBSPKKKNOLKVLHLSRQK
KIRQLRSQCATWKVICKSISQFPGINLDLGSQVKVLIIPKEBCKMPAGEBQPV
"
BASE COUNT 224 a 170 c 216 g 130 t
ORIGIN
Query Match 30.5%; Score 204; DB 9; Length 740;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
Qy 319 GTCCAGAGAACCCATGATGAGAGAGCTTAAGAAGAGAAACCCATTAAGTCGAA 378
Db 383 GTCCAGAGAGCCAGTAATGGAGAGCCCAAAAAGAGAACCCAGAGTCGAAAGTCGGGA 442
Qy 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGCTGAGATTCAGATTCAGTGCCTGA 438
Db 443 TCCTACACCTGGCAGCAGCAGAGAGATCAGATCAGTGAATCCAGTCGCGGA 502
Qy 439 CTGGAAGCCGATCTCCAGAGACTATGTCAGACAAAGACTGGGGATGGATGTGAAGGT-G 497
Db 503 CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGG 562
Qy 498 GTACTGATCTCAAGGGAGATCTACCAAAAGCAGAGCACTTTAAATGCCAGAGCAG 557
Db 563 GTTCGGCTCAAGGTGAAGATAATACCTTAAGAGAACACTGTAAATGCCAGAGCAG 622
Qy 558 GTGAAGGGAATCACAGGTTTAAAGGAAGATAAGCTGAAACACAACTGTTTTTATA 617
Db 623 GTGAGAGCAACCAAGTTTAAATGAGACAGCTGAAACCAACGCAAGCTGGTTTTATA 682
Qy 618 TTAGATATTACTTTAAGAGCTCTTAATAAATTT 653
Db 683 TTAGATATTGACTTAAACTATCTCAATAAAGTTT 718
RESULT 30
BD109661 767 bp DNA linear PAT 18-SEP-2002
LOCUS EST and encoded human protein.
DEFINITION BD109661
ACCESSION BD109661.1 GI:23204479
VERSION JP 2002010789-A/1738.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 767)
AUTHORS Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
TITLE EST and encoded human protein
JOURNAL Patent: JP 2002010789-A 1738 15-JAN-2002;
GENSET CORP
COMMENT OS Homo sapiens (human)
PN JP 2002010789-A/1738
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT,JEAN' EYE PI
GIORDANO
PC C12N15/09,C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19, PC
C12N1/21,
PC C12N5/10,C12P21/02,C12P21/08,C12Q1/68,C12N15/00,C12N5/00, PC
C12N15/00
CC Von Heijne matrix
CC score 4.6999980926514
CC seq LCALLSNFCPSST/VK
FH Key Location/Qualifiers
FT CDS 83..658

FT sig_peptide 83..157.
Location/Qualifiers
source 1..767
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 222 a 193 c 187 g 165 t
ORIGIN
Query Match 30.5%; Score 204; DB 6; Length 767;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
Qy 319 GTCCAGAGAACCCATGATGAGAGAGCTTAAGAAGAGAAACCCATTAAGTCGAA 378
Db 399 GTCCAGAGAGCCAGTAATGGAGAGCCCAAAAAGAGAACCCAGAGTCGAAAGTCGGGA 458
Qy 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGCTGAGATTCAGATTCAGTGCCTGA 438
Db 459 TCCTACACCTGGCAGCAGCAGAGAGATCAGGATACAGCTGAGATCCAGTCGCGGA 518
Qy 439 CTGGAAGCCGATCTCCAGAGACTATGTCAGACAAAGACTGGGGATGGATGTGAAGGT-G 497
Db 519 CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGATTGG 578
Qy 498 GTACTGATGTCAGGGAGAGATCTTACCAAAAGCAGAGCACTTTAAATGCCAGAGCAG 557
Db 579 GTTCGGCTCAAGGTGAAGATAATACCTTAAGAGAGAACACTGTAAATGCCAGAGCAG 638
Qy 558 GTGAAGGGAATCACAGGTTTAAAGGAAGATAAGCTGAAACACAACTGTTTTTATA 617
Db 639 GTGAGAGCAACCAAGTTTAAATGAGAGCAAGCTGAAACACGCAAGCTGGTTTTATA 698
Qy 618 TTAGATATTACTTTAAGAGCTCTTAATAAATTT 653
Db 699 TTAGATATTGACTTAAACTATCTCAATAAAGTTT 734
RESULT 31
AR273067 781 bp DNA linear PAT 10-APR-2003
LOCUS Sequence 808 from patent US 6504010.
DEFINITION AR273067
ACCESSION AR273067
VERSION AR273067.1 GI:29704952
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 781)
AUTHORS Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S.,
Carter,D., Retter,M.W., Mannion,J. and Fan,L.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: US 6504010-A 808 07-JAN-2003;
FEATURES Location/Qualifiers
source 1..781
/organism="unknown"
BASE COUNT 230 a 181 c 234 g 136 t
ORIGIN
Query Match 30.5%; Score 204; DB 6; Length 781;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
Qy 319 GTCCAGAGAACCCATGATGAGAGAGCTTAAGAAGAGAAACCCATTAAGTCGAA 378
Db 422 GTCCAGAGAGCCAGTAATGGAGAGCCCAAAAAGAGAACCCAGAGTCGAAAGTCGGGA 481
Qy 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGCTGAGATTCAGATTCAGTGCCTGA 438
Db 482 TCCTACACCTGGCAGCAGCAGAGAGATATCGATACAGCTGAGATCCAGTCGCGGA 541
Qy 439 CTGGAAGCCGATCTCCAGAGACTATGTCAGACAAAGACTGGGGATGGATGTGAAGGT-G 497

Db 542 CATGAGAGTGTCTGCAAGAGCTGCATCAGTCAAAACACCGGGATAAATCTGATTGG 601
QY 498 GTACTGATGTAAGGGAGATCTTACCAAGACGAGACACTTTAAATGCCAGAGCAG 557
Db 602 GTTCGGCGTCAAGGTGAAGATAACCTAAAGAGAACACTGTAAATGCCAGAGCAG 661
QY 558 GTGAAGGGAAATCACAGGTTTAAAGGAAGATAAGCTGAACACACAACTGTTTTATA 617
Db 662 GTGAAGAGCACACCAAGTTTAAATGAAGCAAGCTGAACACACGAGCTGGTTTTATA 721
QY 618 TTAGATATTCTTAAAGAGCTTTAAATATTTT 653
Db 722 TTAGATATTCTTAAAGAGCTTTAAATATTTT 757

RESULT 32
AR276648 781 bp DNA linear PAT 10-APR-2003
LOCUS
DEFINITION Sequence 808 from patent US 6509448.
ACCESSION AR276648
VERSION AR276648.1 GI:29710295
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 781)
AUTHORS Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S., Carter,D., Retter,M.W., Mannion,J., Fan,L. and Wang,A.
TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: US 6509448-A 808 21-JAN-2003;
FEATURES Location/Qualifiers
source 1..781
BASE COUNT 230 a 181 c 234 g 136 t
ORIGIN
Query Match 30.5%; Score 204; DB 6; Length 781;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGAGAACCCACTGATGAAGAGCCTAAAGAGAGAACCCACCTAAAGTCCGAA 378
Db 422 GTCCAGAGAACCCACTGATGAAGAGCCTAAAGAGAGAACCCACCTAAAGTCCGAA 481
QY 379 TCCTACACCTGATGATGAAGAGCCTAAAGAGAGAACCCACCTAAAGTCCGAA 438
Db 482 TCCTACACCTGATGATGAAGAGCCTAAAGAGAGAACCCACCTAAAGTCCGAA 541
QY 439 CTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGAGCTGGGATGATGAGT-G 497
Db 542 CATGGAAGTGTCTGCAAGAGCTGCATCAGTCAAAACACCGGGATAAATCTGATTGG 601
QY 498 GTACTGATGTCAGGGGAGAGATTCTACCAAGACGAGACCTTTAAATGCCAGAGCAG 557
Db 602 GTTCGGCGTCAAGGTGAAGATAACCTAAAGAGAACACTGTAAATGCCAGAGCAG 661
QY 558 GTGAAGGGAAATCACAGGTTTAAAGGAAGATAAGCTGAACACACAACTGTTTTATA 617
Db 662 GTGAAGAGCACACCAAGTTTAAATGAAGCAAGCTGAACACACGAGCTGGTTTTATA 721
QY 618 TTAGATATTCTTAAAGAGCTTTAAATATTTT 653
Db 722 TTAGATATTCTTAAAGAGCTTTAAATATTTT 757

RESULT 33
AX063181 781 bp DNA linear PAT 24-JAN-2001
LOCUS
DEFINITION Sequence 808 from Patent WO0100828.
ACCESSION AX063181
VERSION AX063181.1 GI:12541027

KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S., Carter,D., Retter,M.W. and Mannion,J.
TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: WO 0100828-A 808 04-JAN-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1..781
BASE COUNT 230 a 181 c 234 g 136 t
ORIGIN
Query Match 30.5%; Score 204; DB 6; Length 781;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGAGAACCCACTGATGAAGAGCCTAAAGAGAGAACCCACCTAAAGTCCGAA 378
Db 422 GTCCAGAGAACCCACTGATGAAGAGCCTAAAGAGAGAACCCACCTAAAGTCCGAA 481
QY 379 TCCTACACCTGATGATGAAGAGCCTAAAGAGAGAACCCACCTAAAGTCCGAA 438
Db 482 TCCTACACCTGATGATGAAGAGCCTAAAGAGAGAACCCACCTAAAGTCCGAA 541
QY 439 CTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGAGCTGGGATGATGAGT-G 497
Db 542 CATGGAAGTGTCTGCAAGAGCTGCATCAGTCAAAACACCGGGATAAATCTGATTGG 601
QY 498 GTACTGATGTCAGGGGAGAGATTCTACCAAGACGAGACCTTTAAATGCCAGAGCAG 557
Db 602 GTTCGGCGTCAAGGTGAAGATAACCTAAAGAGAACACTGTAAATGCCAGAGCAG 661
QY 558 GTGAAGGGAAATCACAGGTTTAAAGGAAGATAAGCTGAACACACAACTGTTTTATA 617
Db 662 GTGAAGAGCACACCAAGTTTAAATGAAGCAAGCTGAACACACGAGCTGGTTTTATA 721
QY 618 TTAGATATTCTTAAAGAGCTTTAAATATTTT 653
Db 722 TTAGATATTCTTAAAGAGCTTTAAATATTTT 757

RESULT 34
AX368098 781 bp DNA linear PAT 16-FEB-2002
LOCUS
DEFINITION Sequence 808 from Patent WO0204514.
ACCESSION AX368098
VERSION AX368098.1 GI:18856173
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Wang,T., Watanabe,Y., Henderson,R.A., Johnson,J.C., Retter,M.W., Marnerakis,M., Carter,D., Fanger,G.R., Vedvick,T.S., Bangur,C.S., McNabb,A., Fanger,N., Switzer,A., McNeill,P.D. and Clapper,J.D.
TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: WO 0204514-A 808 17-JAN-2002;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1..781
BASE COUNT 230 a 181 c 234 g 136 t
ORIGIN
Query Match 30.5%; Score 204; DB 6; Length 781;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGAGAACCCACTGATGAAGAGCCTAAAGAGAGAACCCACCTAAAGTCCGAA 378
Db 422 GTCCAGAGAACCCACTGATGAAGAGCCTAAAGAGAGAACCCACCTAAAGTCCGAA 481
QY 379 TCCTACACCTGATGATGAAGAGCCTAAAGAGAGAACCCACCTAAAGTCCGAA 438
Db 482 TCCTACACCTGATGATGAAGAGCCTAAAGAGAGAACCCACCTAAAGTCCGAA 541
QY 439 CTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGAGCTGGGATGATGAGT-G 497
Db 542 CATGGAAGTGTCTGCAAGAGCTGCATCAGTCAAAACACCGGGATAAATCTGATTGG 601
QY 498 GTACTGATGTCAGGGGAGAGATTCTACCAAGACGAGACCTTTAAATGCCAGAGCAG 557
Db 602 GTTCGGCGTCAAGGTGAAGATAACCTAAAGAGAACACTGTAAATGCCAGAGCAG 661
QY 558 GTGAAGGGAAATCACAGGTTTAAAGGAAGATAAGCTGAACACACAACTGTTTTATA 617
Db 662 GTGAAGAGCACACCAAGTTTAAATGAAGCAAGCTGAACACACGAGCTGGTTTTATA 721
QY 618 TTAGATATTCTTAAAGAGCTTTAAATATTTT 653
Db 722 TTAGATATTCTTAAAGAGCTTTAAATATTTT 757


```
BASE COUNT      230 a      181 c      234 g      136 t
ORIGIN
Query Match      30.5%; Score 204; DB 6; Length 781;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGGAGCCACTGATGAGAGCCCTTAAGAAGAGAACACCCACTAAAGTCGAA 378
Db 422 GTCCAGGAGCCAGTAATGGAGGCCCCAAAAGAAGAACAGCAGCTGAAGTCGGGA 481
QY 379 TCTTACACCTGATCAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGA 438
Db 482 TCTTACACCTGGCAGCAGACAGAGAGATCAGGATCAGCTGAGATCCAGTCGCGA 541
QY 439 CTTGGAGCCGATCTCCAGGACTATGTCAGACRAAGCTGGGATGATGAGGT-G 497
Db 542 CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAACACCGGGGATAAATCTGGATTGG 601
QY 498 GTACTGATGTCAAGGGGAAGATTCTACCAAAAGCAGACACTTTAAATGCCAGAGCAG 557
Db 602 GTTCGGCGTCAAGGTGAAGATAATACCTTAAGAGGAACACACTGTAATGCCAGAGCAG 661
QY 558 GTGAGGGAATACAGGTTTAAAGAGATTAAGCTGAACACACAACTCTTTTATA 617
Db 662 GTGAGAGCAACCAAGTTTAAATGAAGCAAGCTGAACCAACCAAGCTGGTTTATA 721
QY 618 TTAGATATTTTACTTTAAAGAGCTCTTAATAATTTT 653
Db 722 TTAGATATTTGACTTTAACTTCTCTCAATAAGTTT 757
RESULT 35
LOCUS AR275665 528 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 1 from patent US 6509172.
ACCESSION AR275665
VERSION AR275665.1 GI:29709167
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 528)
AUTHORS De Backer, O., Van den Eynde, B. and Boon-Falleur, T.
TITLE Isolated, truncated nucleic acid which are members of the gage, and
uses thereof
JOURNAL Patent: US 6509172-A 1 21-JAN-2003;
FEATURES
source
1. 528
Location/Qualifiers
BASE COUNT 158 a 112 c 151 g 107 t
ORIGIN
Query Match 27.3%; Score 182.8; DB 6; Length 528;
Best Local Similarity 67.6%; Pred. No. 9.6e-40;
Matches 288; Conservative 0; Mismatches 132; Indels 6; Gaps 2;
QY 217 GTGAATATGAGTTGGCGAGAGATCAATATAGCGCTAGCCCAAGAAGATTACA 276
Db 66 GTGAATATGAGTTGGCGAGAGATCAATATAGCGCTAGCCCAAGAAGATTACA 125
QY 277 GCCTCTGAGCTGATTGGGCTATGCTTACTGGCTCCCTTTGTCCAGAGAACCCACTGA 336
Db 126 GCCTCTGAAATGATTGGGCTATGCGGCC---CGAGCAGTTTCAAGTGAAGTGAAC 182
QY 337 TGAAGAGCTTAAGAGAGAGAACCCACTAAAGTCGGAATCTTACACCTGATCAGAA 396
Db 183 AGCAACACCTTGAAGAGAGAGAACCCACTAAAGTCGGAATCTTACACCTGATCAGAA 242
QY 397 GAGAGAGATGATCAGGCTGATGAGATTCAAGTGCCTGAGCTTCCAGAGCCGATCTCCA 456
Db 243 GGGAGAGATG---AGGGAGCATCTGCGAGTCAAGGGCCGAGAGCTTGAAGCTGATAGCA 299
```

```
QY 457 GGAGCTATGTCAGACAAAGACTGGGATGGATGCTGAAGCTGGTACTGATGTCAAGGGAA 516
Db 300 GGAACAGGGTCAACCCACAGACTGGGTGAGTGTGAGTGTGATGTGCTTCTGATGGCAGAGAT 359
QY 517 GATTTCACAAAAGCAGACACTTTTAAATGCCAGAGCAGGTGAAGGGAATCAAGGT 576
Db 360 GGACCCGCCAAATCCAGAGAGGTGAAAACGCCCTGAAGAGGTGAAAAGCAATCACAGTG 419
QY 577 TTAAGAGAGATTAAGCTGAAACACAAACACTGTTTATATATTAGATATTACTTTAAA 636
Db 420 TTAAGAGAGACACAGCTTTGAATGATGAGGCTCTCTTATGTTGAAATTTGTTCAATTAA 479
QY 637 GAGTCT 642
Db 480 AATTCT 485
RESULT 36
LOCUS AF055473 528 bp mRNA linear PRI 01-MAY-2000
DEFINITION Homo sapiens GAGE-8 mRNA, complete cds.
ACCESSION AF055473
VERSION AF055473.1 GI:3511022
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 528)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
De Backer, O., Arden, K.C., Boretti, M., Vantomme, V., De Smet, C.,
Czekay, S., Viars, C.S., De Plaen, E., Brasseur, F., Chomez, P., Van den
Eynde, B., Boon, T. and van der Bruggen, P.
TITLE Characterization of the GAGE genes that are expressed in various
human cancers and in normal testis
JOURNAL Cancer Res. 59 (13), 3157-3165 (1999)
MEDLINE 99323388
PUBMED 10397259
REFERENCE 2 (bases 1 to 528)
AUTHORS De Backer, O.R.Y.
TITLE Direct Submission
JOURNAL Submitted (24-MAR-1998) Ludwig Institute for Cancer Research,
Brussels Branch, 74 av. Hippocrate, Brussels B-1200, Belgium
FEATURES
source
1. 528
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
73..423
/codon_start=1
/product="GAGE-8"
/protein_id="AAC33676.1"
/db_xref="GI:3511023"
/translation="MSWRGSRSTYRPRRYVEPMIGMRPEQFSDVEVPATPESGE
PATQRQDFAAQSGEDGASAGGPKPEADSGQHPQTCEDGPDGQEMDPNPE
EVKTPESGEKQSQ"
BASE COUNT 158 a 112 c 151 g 107 t
ORIGIN
Query Match 27.3%; Score 182.8; DB 9; Length 528;
Best Local Similarity 67.6%; Pred. No. 9.6e-40;
Matches 288; Conservative 0; Mismatches 132; Indels 6; Gaps 2;
QY 217 GTGAATATGAGTTGGCGAGAGATCAATATAGCGCTAGCCCAAGAAGATTACA 276
Db 66 GTGAATATGAGTTGGCGAGAGATCAATATAGCGCTAGCCCAAGAAGATTACA 125
QY 277 GCCTCTGAGCTGATTGGGCTATGCTTACTGGCTCCCTTTGTCCAGAGAACCCACTGA 336
Db 126 GCCTCTGAAATGATTGGGCTATGCGGCC---CGAGCAGTTTCAAGTGAAGTGAAC 182
QY 337 TGAAGAGCTTAAGAGAGAGAACCCACTAAAGTCGGAATCTTACACCTGATCAGAA 396
Db 183 AGCAACACCTTGAAGAGAGAGAACCCACTAAAGTCGGAATCTTACACCTGATCAGAA 242
```

QY 397 GAGAGAAGATGATCAGGCTGACGCTGAGATTCAAGTGTGCTGACCTGGGAAGCGGATCTCCA 456
 |||||
 Db 243 GGGAGAGGATG---AGGAGCATCTGCAGGTCAAGGGCGAAGCTGAAGCTGATAGCCA 299
 |||||
 QY 457 GGAGCTATGTGACAGAAAGCTGGGATGGATGTGAAGTGTACTGTATGTCAAGGGAA 516
 |||||
 Db 300 GGAACAGGCTACCCACAGACTGGGTGTGAGTGTGAAGATGCTCTGTATGGCAGGAT 359
 |||||
 QY 517 GATTCTACCAAGCAGAGACACTTTAAATGCCAGAGCAGGTGAAGGGAATCACAGGT 576
 |||||
 Db 360 GGACCCGCCAATCCAGAGGAGGTGAACCGCTGAGNAGGTGAAGCAATCACAGTG 419
 |||||
 QY 577 TTAAGGAAGATGAAGCTGAAACACACAACTGTTTTATATAGATATTTACTTTAAA 636
 |||||
 Db 420 TTAAGGAAGACACGTTGAATGATGAGCGTCTCTATGTTGGAATTTGTTCAITAA 479
 |||||
 QY 637 GAGTCT 642
 |||||
 Db 480 AATTCT 495
 |||||

RESULT 37
 HSU19143 530 bp mRNA linear PRI 04-DEC-1995
 LOCUS Human GAGE-2 protein mRNA, complete cds.

DEFINITION U19143
 ACCESSION U19143
 VERSION U19143.1 GI:914900

KEYWORDS
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 530)
 AUTHORS Van den Eynde, B., Peeters, O., De Backer, O., Gaugler, B., Lucas, S. and Boon, T.

TITLE A new family of genes coding for an antigen recognized by

JOURNAL autologous cytolytic T lymphocytes on a human melanoma

MEDLINE J. Exp. Med. 182 (3), 689-698 (1995);

PUBMED 95378788

REFERENCE 2 (bases 1 to 530)

AUTHORS Van Den Eynde, B.J.

TITLE Direct Submission

JOURNAL Submitted (28-DEC-1994) Benoit J Van Den Eynde, Ludwig Institute

For Cancer Research, 74 Avenue Hippocrate, BRUSSELS, 1200, BELGIUM

FEATURES
 Location/Qualifiers

1..530

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/sex="female"

/cell_line="M2-MEL.43"

/tissue_type="melanoma"

/dev_stage="adult"

84..434

/codon_start=1

/product="GAGE-2 protein"

/protein_id="AAA82745.1"

/db_xref="GI:914901"

/translation="MSWRGRSTYRPRRYVEPPMIGPMRPEQSFDEVPATPEEGE

PATORDPAAQEGEDBGAGGQPKPEAHSRQGHPTGCECEDPGDQMDPPNPE

EVTPEGEKQSC"

BASE COUNT 152 a 116 c 155 g 107 t

ORIGIN

Query Match 27.1%; Score 181.2; DB 9; Length 530;

Best Local Similarity 67.4%; Pred. No. 2.7e-39;

Matches 287; Conservative 0; Mismatches 133; Indels 6; Gaps 2;

QY 217 GTGAATATGATGTTGGCGAGGAGATCAACATATAGCGCTAGGCCAAGAGAGTTTACA 276

|||||

Db 77 GTGAATATGATGTTGGCGAGGAGATCGACCTATCGCGCTAGACCAAGACGCTAGTAGA 136

|||||

QY 277 GCCTCTCAGCTGATGGGGGTATGCTTACTGTGGTCCCTTTGTGCCAGAACCCACTGA 336

|||||

Db 137 GCCTCTCAGCTGATGGGGGTATGCTTACTGTGGTCCCTTTGTGCCAGAACCCACTGA 193

|||||

QY 337 TGAAGCGCTTAAGAAGAGAACCCACCACCTAAAGTCGGAAATCCTACCTGATCAGAA 396

|||||

Db 194 AGCAACCTTAAGAAGAGAACCCACCACCTAAAGTCGGAAATCCTACCTGATCAGAA 253

|||||

QY 397 GAGAGAAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAAGCCGATCTCCA 456

|||||

Db 254 GGGAGAGGATG---AGGAGCATCTGCAGGTCAAGGGCCGAAAGCCTGAAGCTCATAGCCA 310

|||||

QY 457 GGAGCTATGTGACAGCAAAAGACTGGGGTGGATGTGAAGTGGTACTGATGTCAGGGGAA 516

|||||

QY 277 GCCTCTCAGCTGATGGGGGTATGCTTACTGTGGTCCCTTTGTGCCAGAACCCACTGA 336
 |||||
 Db 137 GCCTCTCAGCTGATGGGGGTATGCTTACTGTGGTCCCTTTGTGCCAGAACCCACTGA 193
 |||||
 QY 337 TGAAGCGCTTAAGAAGAGAACCCACCACCTAAAGTCGGAAATCCTACCTGATCAGAA 396
 |||||
 Db 194 AGCAACCTTAAGAAGAGAACCCACCACCTAAAGTCGGAAATCCTACCTGATCAGAA 253
 |||||
 QY 397 GAGAGAAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAAGCCGATCTCCA 456
 |||||
 Db 254 GGGAGAGGATG---AGGAGCATCTGCAGGTCAAGGGCCGAAAGCCTGAAGCTCATAGCCA 310
 |||||
 QY 457 GGAGCTATGTGACAGCAAAAGACTGGGGTGGATGGATGTGAAGTGTACTGATCTCAAGGGGAA 516
 |||||
 Db 311 GGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGTGTGCTTCTGATGGCAGAGAT 370
 |||||
 QY 517 GATTCTACCAAAAGCAGAGCACTTTAAATGCCAGAGCAGGTGAAGGGAATCACAGGT 576
 |||||
 Db 371 GGACCCGCCAATCCAGAGGAGGTGAACCGCTGAGNAGGTGAAGCAATCACAGTG 430
 |||||
 QY 577 TTAAGGAAGATGAAGCTGAAACACACAACTGTTTTATATAGATATTTACTTTAAA 636
 |||||
 Db 431 TTAAGGAAGACACGTTGAATGATGAGCGTCTCTATGTTGGAATTTGTTCAITAA 490
 |||||
 QY 637 GAGTCT 642
 |||||
 Db 491 AATTCT 496
 |||||

RESULT 38

AR028488

LOCUS

DEFINITION Sequence 14 from patent US 5858689.

ACCESSION AR028488

VERSION AR028488.1 GI:5940461

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 538)

AUTHORS van der Bruggen, P., van den Eynde, B., DeBacker, O. and

Boon-Pallieur, T.

TITLE Isolated peptides derived from the gage tumor rejection antigen

JOURNAL precursor and uses thereof

Patent: US 5858689-A 14 12-JAN-1999;

FEATURES
 Location/Qualifiers

1..538

/organism="unknown"

BASE COUNT 160 a 116 c 155 g 107 t

ORIGIN

Query Match 27.1%; Score 181.2; DB 6; Length 538;

Best Local Similarity 67.4%; Pred. No. 2.7e-39;

Matches 287; Conservative 0; Mismatches 133; Indels 6; Gaps 2;

QY 217 GTGAATATGATGTTGGCGAGGAGATCAACATATAGCGCTAGGCCAAGAGAGTTTACA 276

|||||

Db 77 GTGAATATGATGTTGGCGAGGAGATCGACCTATCGCGCTAGACCAAGACGCTAGTAGA 136

|||||

QY 277 GCCTCTCAGCTGATGGGGGTATGCTTACTGTGGTCCCTTTGTGCCAGAACCCACTGA 336

|||||

Db 137 GCCTCTCAGCTGATGGGGGTATGCTTACTGTGGTCCCTTTGTGCCAGAACCCACTGA 193

|||||

QY 337 TGAAGCGCTTAAGAAGAGAACCCACCACCTAAAGTCGGAAATCCTACCTGATCAGAA 396

|||||

Db 194 AGCAACCTTAAGAAGAGAACCCACCACCTAAAGTCGGAAATCCTACCTGATCAGAA 253

|||||

QY 397 GAGAGAAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAAGCCGATCTCCA 456

|||||

Db 254 GGGAGAGGATG---AGGAGCATCTGCAGGTCAAGGGCCGAAAGCCTGAAGCTCATAGCCA 310

|||||

QY 457 GGAGCTATGTGACAGCAAAAGACTGGGGTGGATGTGAAGTGGTACTGATGTCAGGGGAA 516

|||||

Db 311 GGAAAGGGTCAACACAGCTGGGTGTGAGTGTGAAGATGCTGATGGCAGGAGAT 370
Qy 517 GATTCTACCAAGCAGACACTTTAAATGCGCAGAGCAGGTGAAGGAAATCACAGGT 576
Db 371 GGACCGCCAAATCCAGAGAGGTGAAAACGCGCTGAAGAGGTGAAAAGCAATCACAGT 430
Qy 577 TTAAGGAAGATAGCTGAAACCAACAACTGTTTTATATTAGATATTACTTTAA 636
Db 431 TTAAGGAAGACACCTTGAATGATCGAGGTGCTCTATGTTGGAATTTGTTCAATA 490
Qy 637 GAGTCT 642
Db 491 AATTCT 496

RESULT 39
BD132467
LOCUS 538 bp DNA linear PAT 18-SEP-2002
DEFINITION Isolated, nucleic acid molecules which code for GAGE tumor
rejection antigen, the tumor rejection antigen, and uses thereof.
ACCESSION BD132467
VERSION BD132467.1 GI:23227412
KEYWORDS JP 2002507112-A/7.
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 538)
AUTHORS DeBacker, O., Bynde, B.V.D. and Falleur, T.B.
TITLE Isolated, nucleic acid molecules which code for GAGE tumor
rejection antigen, the tumor rejection antigen, and uses thereof
JOURNAL Patent: JP 2002507112-A 7 05-MAR-2002;
LUDWIG INSTITUTE FOR CANCER RESEARCH
COMMENT PN JP 2002507112-A/7
PD 05-MAR-2002
PP 23-JUN-1997 JP 1998503430
PR 24-JUN-1996 US 08/689161
PI OLIVIER DEBACKER, BENOIT VAN DEN EYNDE, THIERRY BOON FALLEUR PC.
A61K38/00, A61K45/05, C07K7/00, C07K14/82, C12N15/00 CC
Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.

FEATURES
source 1..538
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT 160 a 116 c 155 g 107 t
ORIGIN

Query Match 27.1%; Score 181.2; DB 6; Length 538;
Best Local Similarity 67.4%; Pred. No. 2.7e-39;
Matches 287; Conservative 0; Mismatches 133; Indels 6; Gaps 2;

Qy 217 GTGAATATGAGTTGGCGAGAGATCAACATATAGCCCTAGGCCAAGAGAGTTTACA 276
Db 77 GTGAATATGAGTTGGCGAGAGATCGACCTATCGCCCTAGACCAAGAGCTACGTAGA 136

Qy 277 GCCTCTGAGCTGATTTGGGCTATGCTTACTGGCTCCCTTTGTCCTCAGGACCCACTGA 336
Db 137 GCCTCTGAGATGATTTGGGCTATGCGGCC---CGAGAGTTGAGTGATGAGTGGGAACC 193

Qy 337 TGAAGAGCCTAAAGAGAGAAACCCACCTAAAGTCGGAATCCTACCTGATCAGTACGAA 396
Db 194 AGCAACACCTGAAGAGAGGGAACCCAGCACTCAAGCTCAGGATCCTGCTCAGGA 253

Qy 397 GAGAGAGATGATCAGGCTGAGTGTGAGTCAAGTGCCTGACCTGGAAGCCGATCTCCA 456
Db 254 GGGAGAGATG---AGGAGAGATCTGCAAGTCAAGGGCCGAGCCTGAAGCTCATAGCCA 310

Qy 457 GGAGCTATGTGAGCAAAAGACTGGGGATGGATGTGAAGTGGTACTGATGTCAGGGGAA 516
Db 311 GGAACAGGTCACCCACAGACTGGGTGTGAGTGTGAAGTGGTCTGATGGCAGGAGAT 370

Qy 517 GATTCTACCAAGCAGACACTTTAAATGCGCAGAGCAGGTGAAGGAAATCACAGGT 576
Db 371 GGACCGCCAAATCCAGAGAGGTGAAAACGCGCTGAAGAGGTGAAAAGCAATCACAGT 430

Qy 577 TTAAGGAAGATAGCTGAAACCAACAACTGTTTTATATTAGATATTACTTTAA 636
Db 431 TTAAGGAAGACACCTTGAATGATCGAGGTGCTCTATGTTGGAATTTGTTCAATA 490
Qy 637 GAGTCT 642
Db 491 AATTCT 496

RESULT 41

Qy 517 GATTCTACCAAGCAGACACTTTAAATGCGCAGAGCAGGTGAAGGAAATCACAGGT 576
Db 371 GGACCGCCAAATCCAGAGAGGTGAAAACGCGCTGAAGAGGTGAAAAGCAATCACAGT 430
Qy 577 TTAAGGAAGATAGCTGAAACCAACAACTGTTTTATATTAGATATTACTTTAA 636
Db 431 TTAAGGAAGACACCTTGAATGATCGAGGTGCTCTATGTTGGAATTTGTTCAATA 490
Qy 637 GAGTCT 642
Db 491 AATTCT 496

RESULT 40
BD132467
LOCUS 538 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 14 from patent US 5648226.
ACCESSION BD132467
VERSION BD132467.1 GI:2476645
KEYWORDS .
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 538)
AUTHORS Van den Eynde, B., DeBacker, O. and Boon-Falleur, T.
TITLE Isolated peptides derived from tumor rejection antigens, and their
use
JOURNAL Patent: US 5648226-A 14 15-JUL-1997;
FEATURES Location/Qualifiers
source 1..538
/organism="unknown"
BASE COUNT 160 a 116 c 155 g 107 t
ORIGIN

Query Match 27.1%; Score 181.2; DB 6; Length 538;
Best Local Similarity 67.4%; Pred. No. 2.7e-39;
Matches 287; Conservative 0; Mismatches 133; Indels 6; Gaps 2;

Qy 217 GTGAATATGAGTTGGCGAGAGATCAACATATAGCCCTAGGCCAAGAGAGTTTACA 276
Db 77 GTGAATATGAGTTGGCGAGAGATCGACCTATCGCCCTAGACCAAGAGCTACGTAGA 136

Qy 277 GCCTCTGAGCTGATTTGGGCTATGCTTACTGGCTCCCTTTGTCCTCAGGACCCACTGA 336
Db 137 GCCTCTGAGATGATTTGGGCTATGCGGCC---CGAGAGTTGAGTGATGAGTGGGAACC 193

Qy 337 TGAAGAGCCTAAAGAGAGAAACCCACCTAAAGTCGGAATCCTACCTGATCAGTACGAA 396
Db 194 AGCAACACCTGAAGAGAGGGAACCCAGCACTCAAGCTCAGGATCCTGCTCAGGA 253

Qy 397 GAGAGAGATGATCAGGCTGAGTGTGAGTCAAGTGCCTGACCTGGAAGCCGATCTCCA 456
Db 254 GGGAGAGATG---AGGAGAGATCTGCAAGTCAAGGGCCGAGCCTGAAGCTCATAGCCA 310

Qy 457 GGAGCTATGTGAGCAAAAGACTGGGGATGGATGTGAAGTGGTACTGATGTCAGGGGAA 516
Db 311 GGAACAGGTCACCCACAGACTGGGTGTGAGTGTGAAGTGGTCTGATGGCAGGAGAT 370

Qy 517 GATTCTACCAAGCAGACACTTTAAATGCGCAGAGCAGGTGAAGGAAATCACAGGT 576
Db 371 GGACCGCCAAATCCAGAGAGGTGAAAACGCGCTGAAGAGGTGAAAAGCAATCACAGT 430

Qy 577 TTAAGGAAGATAGCTGAAACCAACAACTGTTTTATATTAGATATTACTTTAA 636
Db 431 TTAAGGAAGACACCTTGAATGATCGAGGTGCTCTATGTTGGAATTTGTTCAATA 490
Qy 637 GAGTCT 642
Db 491 AATTCT 496

RESULT 41

BC018052
LOCUS BC018052 551 bp mRNA linear PRI 06-DEC-2001
DEFINITION Homo sapiens, G antigen 8, clone MGC:26395 IMAGE:4812462, mRNA,
complete cds.
ACCESSION BC018052
VERSION BC018052.1 GI:17390105
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LILNL at: <http://image.llnl.gov>
Series: IRAK Plate: 32 Row: k Column: 1
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4503878.

FEATURES

Location/Qualifiers
1..551
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocusID:26749"
/db_xref="taxon:9606"
/clone="MGC:26395 IMAGE:4812462"
/tissue_type="Brain, hippocampus"
/clone_lib="NIH MGC 95"
/lab_host="DH10B"
/note="Vector: pBluescript"

CDS

97..447
/codon_start=1
/product="G antigen 8"
/protein_id="AAH18052.1"
/db_xref="GI:17390106"
/translations="MSWRGRSTYRPRPRYPPEPMIGRMRPEQFSDEVEPATPEEGE
PATQRQDPAAQGEDEGASAGQGPKEADSOQHPTQGCEDGPDQEMDPNPE
EVTPEGEKQSQ"

BASE COUNT 164 a 118 c 160 g 109 t

ORIGIN

Query Match 27.1%; Score 181.2; DB 9; Length 551;
Best Local Similarity 67.4%; Pred. No. 2.7e-39;
Matches 287; Conservative 0; Mismatches 133; Indels 6; Gaps 2;

QY 217 GTGAATATCAGTTGGGAGAGATCAACATATAGCCCTAGCCCAAGAGAGTTTACA 276
DB 90 GTGAATATCAGTTGGGAGAGATCGACTATCGCCCTAGACCAAGACGCTACGTAGA 149
QY 277 GCCTCCTGAGCTGATGGGCTATGCTTACTGGCTCCCTTTGTCACAGAACCCACTGA 336
DB 150 GCCTCCTGAAATGATTGGGCTATGCGGCC---CGAGCAGTTCAAGTGAAGTGAAC 206
QY 337 TGAAGACCTTAAGAGAGAAACCCACTAAAGTCGGAATCCCTACACCTGATCAGAA 396

DB 207 AGCAACACCTGAAGAAGGGGAACCACTCAACCTCAGGATCCTGCAGCTGCTCAGA 266
QY 397 GAGAGAAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAAGCCGATCTCCA 456
DB 267 GGGAGAGGATG---AGGGAGCATCTCAGGTCAAGGGCCGAAGCCTGAAGCTGATAGCCA 323
QY 457 GGAGCTATGTCAGACAAAGACTGGGGATGGATGTGTGAAGTGGTACTGATGTCAAGGGGAA 516
DB 324 GGAACAGGTCACCCACAGACTGGGTGTGAGTGTGAGATGTCTCTGATGGCAGAGAT 383
QY 517 GATTTCACAAAAGCAGAGACCTTTTAAATGCCAAGCAGGTGAAGGGAATCAGAGT 576
DB 384 GGACCCGCCAAATCCAGAGAGGTGAAACCGCTGAAGAAGGTGAAGAAGCAATCAGAGTG 443
QY 577 TTAAGAGAGATGAAGCTGAACCAACACAACTGTTTTTATATTAGATATTTTACTTTAAA 636
DB 444 TTAAGAAGAGCAGCTTGAATATGATCAGGCTGCTCTTAATGTTGGAATTTGTTCAATA 503
QY 637 GAGTCT 642
DB 504 AATTCT 509
RESULT 42
AR272358 479 bp DNA linear PAT 10-APR-2003
LOCUS Sequence 70 from patent US 6504010.
DEFINITION AR272358
ACCESSION AR272358
VERSION AR272358.1 GI:29704243
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 479)
AUTHORS Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedwick, T.S.,
Carter, D., Retter, M.W., Mannion, J. and Fan, L.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: US 6504010-A 70 07-JAN-2003;
FEATURES Location/Qualifiers
source 1..479
/organism="unknown"
BASE COUNT 163 a 107 c 125 g 84 t
ORIGIN
Query Match 26.9%; Score 180; DB 6; Length 479;
Best Local Similarity 74.0%; Pred. No. 5.8e-39;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGGACCCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCCACTAAAGTCCGAA 378
DB 103 GTCCAGGAGCCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCCACTGAAAGTCGGA 162
QY 379 TCCTACACCTGATCAGAAGAGAGAGATGATCAGGGTGCAGCTGAGATTC----- 428
DB 163 TCCTACACCTGGGAGAGAGAGAGATGATCAGGTACAGCTGAGATCCAGGTGCTGG 222
QY 429 -----AAGTGCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGG 482
DB 223 GAAGGGAATCCGCGACATGGAAGTGTCTGCAAGAGCTGCTCATCAGTCAACACCGGG 282
QY 483 ATGGATGTGAAGGT-GGTACTGATGTCAAGGGGAGAGATTCTACAAAGCAGAGCACTTT 541
DB 283 ATAAATCTGGATTGGGTTCCGGGCTCAAGTGAAGATAATACCTAAAGAGGACACTGT 342
QY 542 AAAATGCCAGAGAGGTGAGGGAATCAGAGTTTAAAGGAGATGAAGCTGAACAC 601
DB 343 AAAATGCCAGAGAGGTGAGGGAATCAGAGTTTAAAGGAGATGAAGCTGAACAC 402
QY 602 ACAAACTGTTTTTATATTAGATATTTTACTTTAAAGAGTCTTAATAAATTTTG 655
DB 403 GCAAGCTGGTTTTATATTAGGATATTTGACTTAAACTATCTCAATAAAGTTTG 456

```
RESULT 43
AR275939
LOCUS AR275939 479 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 70 from patent US 6509448.
ACCESSION AR275939
VERSION AR275939.1 GI:29709584
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE
AUTHORS Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S.,
Carter, D., Retter, M.W., Mannion, J., Fan, B. and Wang, A.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: US 6509448-A 70 21-JAN-2003;
FEATURES
LOCATION/Qualifiers
SOURCE 1. 479
BASE COUNT 163 a 107 c 125 g 84 t
ORIGIN
Query Match 26.9%; Score 180; DB 6; Length 479;
Best Local Similarity 74.0%; Pred. No. 5.8e-39;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGGAAACCCACTGATGAAGAGCCTTAAGAGAAACCCACCCTAAAGTCGGAA 378
Db 103 GTCCAGGAGCCAGTAATGGAGAGCCCAAAAGAGAACCCAGCAGCTGAAGTCGGGA 162
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCGAGCTGAGATTTC----- 428
Db 163 TCCTACACCTGGGCAGCAGACAGAGAGATCAGATACAGCTGAGATCCAGGTGCTGG 222
QY 429 -----AAGTGCTGACCTGGAGCGGATCTCCAGAGCTATGTCCAGAAAGAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGG 282
QY 483 ATGATGTGAAGT-CGTACTGATGTCAGGGGAGAGATTCTACCAAAAGCAGACACTTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGCGTCAGGTGAAGATAATACCTAAAGAGGAACTGT 342
QY 542 AAAATGCCAGAGCAGGTGAAGGGAATCACAGGTTTAAAGGAGATAAGCTGAACAAC 601
Db 343 AAAATGCCAGAGCAGGTGAAGAGCAACCAACAAGTTTAAATGAAGCAAGCTGAACAAC 402
QY 602 ACAACTGTTTTATATTAGATATTTTACTTTAAAGAGCTTTAATAAATTTTG 655
Db 403 GCAAGCTGGTTTTATATTAGGATATTGACTTAAACTATCTCAATAAAGTTTG 456

RESULT 44
AX062443
LOCUS AX062443 479 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 70 from Patent WO0100828.
ACCESSION AX062443
VERSION AX062443.1 GI:12540318
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S.,
Carter, D., Retter, M.W. and Mannion, J.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0100828-A 70 04-JAN-2001;
FEATURES
LOCATION/Qualifiers
SOURCE 1. 479
BASE COUNT 163 a 107 c 125 g 84 t
ORIGIN
Query Match 26.9%; Score 180; DB 6; Length 479;
Best Local Similarity 74.0%; Pred. No. 5.8e-39;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGGAAACCCACTGATGAAGAGCCTTAAGAGAAACCCACCCTAAAGTCGGAA 378
Db 103 GTCCAGGAGCCAGTAATGGAGAGCCCAAAAGAGAACCCAGCAGCTGAAGTCGGGA 162
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCGAGCTGAGATTTC----- 428
Db 163 TCCTACACCTGGGCAGCAGACAGAGAGATCAGATACAGCTGAGATCCCGAGTGTGG 222
QY 429 -----AAGTGCTGACCTGGAGCGGATCTCCAGAGCTATGTCCAGAAAGAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGG 282
QY 483 ATGATGTGAAGT-CGTACTGATGTCAGGGGAGAGATTCTACCAAAAGCAGACACTTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGCGTCAGGTGAAGATAATACCTAAAGAGGAACTGT 342
QY 542 AAAATGCCAGAGCAGGTGAAGGGAATCACAGGTTTAAAGGAGATAAGCTGAACAAC 601
Db 343 AAAATGCCAGAGCAGGTGAAGAGCAACCAACAAGTTTAAATGAAGCAAGCTGAACAAC 402
QY 602 ACAACTGTTTTATATTAGATATTTTACTTTAAAGAGCTTTAATAAATTTTG 655
Db 403 GCAAGCTGGTTTTATATTAGGATATTGACTTAAACTATCTCAATAAAGTTTG 456

/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 163 a 107 c 125 g 84 t
ORIGIN
Query Match 26.9%; Score 180; DB 6; Length 479;
Best Local Similarity 74.0%; Pred. No. 5.8e-39;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGGAAACCCACTGATGAAGAGCCTTAAGAGAAACCCACCCTAAAGTCGGAA 378
Db 103 GTCCAGGAGCCAGTAATGGAGAGCCCAAAAGAGAACCCAGCAGCTGAAGTCGGGA 162
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCGAGCTGAGATTTC----- 428
Db 163 TCCTACACCTGGGCAGCAGACAGAGAGATCAGATACAGCTGAGATCCAGGTGCTGG 222
QY 429 -----AAGTGCTGACCTGGAGCGGATCTCCAGAGCTATGTCCAGAAAGAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGG 282
QY 483 ATGATGTGAAGT-CGTACTGATGTCAGGGGAGAGATTCTACCAAAAGCAGACACTTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGCGTCAGGTGAAGATAATACCTAAAGAGGAACTGT 342
QY 542 AAAATGCCAGAGCAGGTGAAGGGAATCACAGGTTTAAAGGAGATAAGCTGAACAAC 601
Db 343 AAAATGCCAGAGCAGGTGAAGAGCAACCAACAAGTTTAAATGAAGCAAGCTGAACAAC 402
QY 602 ACAACTGTTTTATATTAGATATTTTACTTTAAAGAGCTTTAATAAATTTTG 655
Db 403 GCAAGCTGGTTTTATATTAGGATATTGACTTAAACTATCTCAATAAAGTTTG 456

RESULT 45
AX367360
LOCUS AX367360 479 bp DNA linear PAT 16-FEB-2002
DEFINITION Sequence 70 from Patent WO0204514.
ACCESSION AX367360
VERSION AX367360.1 GI:18855464
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS Wang, T., Watanabe, Y., Henderson, R.A., Johnson, J.C., Retter, M.W.,
Marnezakis, M., Carter, D., Fanger, G.R., Vedvick, T.S., Bangur, C.S.,
McNabb, A., Fanger, N., Switzer, A., McNeill, P.D. and Clapper, J.D.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0204514-A 70 17-JAN-2002;
FEATURES
LOCATION/Qualifiers
SOURCE 1. 479
BASE COUNT 163 a 107 c 125 g 84 t
ORIGIN
Query Match 26.9%; Score 180; DB 6; Length 479;
Best Local Similarity 74.0%; Pred. No. 5.8e-39;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGGAAACCCACTGATGAAGAGCCTTAAGAGAAACCCACCCTAAAGTCGGAA 378
Db 103 GTCCAGGAGCCAGTAATGGAGAGCCCAAAAGAGAACCCAGCAGCTGAAGTCGGGA 162
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCGAGCTGAGATTTC----- 428
Db 163 TCCTACACCTGGGCAGCAGACAGAGAGATCAGATACAGCTGAGATCCCGAGTGTGG 222
```

```
QY 429 -----AAGTGCTGACCTGGAAGCCGATCTCCAGAGCTATGTGAGACAAAGACTGGGG 482
Db 223 GAAGGGAATGCGCGACATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCCGGGG 282
QY 483 ATGGATGTGAAGT-GGTACTGATGTCAAGGGGAAGATTCTACCAAAAGCAGAGCACTTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGCGTCAAGGTGAAGATAATACCTAAAGAGGAACACTGT 342
QY 542 AAAATGCCAGAACGAGTGAAGGAAATCAACAGGTTTAAAGGAAGATAAGCTGAACAAAC 601
Db 343 AAAATGCCAGAACGAGTGAAGGAAATCAACAGGTTTAAAGGAAGATAAGCTGAACAAAC 402
QY 602 ACAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAATAAATTTTG 655
Db 403 GCAAGCTGGTTTTATATTAGATATTTGACTTAAACTATCTCAATAAAGTTTG 456
```

Search completed: January 18, 2004, 12:21:59
Job time : 1695.16 secs

OM nucleic - nucleic search, using sw model

Run on: January 17, 2004, 17:24:54 ; Search time 124.524 Seconds
(without alignments)
14502.633 Million cell updates/sec

Title: US-10-051-835-1
Perfect score: 669
Sequence: 1 cacacagcagcagcagctt.....tttttggcagtcgatctc 669

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :				N_Geneseq 19Jun03:*	
1:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*				
2:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*				
3:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*				
4:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*				
5:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*				
6:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*				
7:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*				
8:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*				
9:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*				
10:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*				
11:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*				
12:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*				
13:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*				
14:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*				
15:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*				
16:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*				
17:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*				
18:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*				
19:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*				
20:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*				
21:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*				
22:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*				
23:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*				
24:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*				
25:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*				

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	581.8	87.0	673	22	AAK51918 Human polynucleoti
2	578.6	86.5	642	24	ASA93876 Human G protein-co
3	558	83.4	756	25	ABT15738 Human cancer/testi
4	549	82.1	580	22	AAFP59637 Human cell cycle a
5	507.8	75.9	611	22	ARD14983 Human NOV4 DNA H
6	469.2	70.1	665	22	AAK52302 Human polynucleoti
7	317	47.4	659	22	AAI58744 Human polynucleoti
8	296.2	44.3	532	22	AAI60530 Human polynucleoti

9	290.2	43.4	509	25	ABX77605 Differentially exp
10	288.6	43.1	520	25	ABT15737 Human cancer/testi
11	282.8	42.0	503	23	AAAS69484 DNA encoding novel
12	242.4	36.2	661	23	AAAS69486 DNA encoding novel
13	215	32.1	320	22	AAH83260 Human ovarian tumo
14	214	32.0	214	22	AAAS24637 Human ovorian PCR-
15	204	30.5	399	22	AAAF68151 Human lung tumour
16	204	30.5	399	22	AAAF68151 Human lung tumour
17	204	30.5	399	22	AAAF68151 Human lung tumour
18	204	30.5	399	22	AAAF68151 Human lung tumour
19	204	30.5	399	22	AAAF68151 Human lung tumour
20	204	30.5	399	22	AAAF68151 Human lung tumour
21	204	30.5	399	22	AAAF68151 Human lung tumour
22	204	30.5	399	22	AAAF68151 Human lung tumour
23	204	30.5	399	22	AAAF68151 Human lung tumour
24	204	30.5	399	22	AAAF68151 Human lung tumour
25	204	30.5	399	22	AAAF68151 Human lung tumour
26	204	30.5	399	22	AAAF68151 Human lung tumour
27	204	30.5	399	22	AAAF68151 Human lung tumour
28	204	30.5	399	22	AAAF68151 Human lung tumour
29	204	30.5	399	22	AAAF68151 Human lung tumour
30	204	30.5	399	22	AAAF68151 Human lung tumour
31	204	30.5	399	22	AAAF68151 Human lung tumour
32	195	29.1	822	22	AAAF68151 Human lung tumour
33	191.2	28.6	618	22	AAAF68151 Human lung tumour
34	183.8	27.5	750	23	AAAS91235 DNA encoding novel
35	181.2	27.1	530	25	ABZ20463 GAGE-2 full length
36	181.2	27.1	535	19	AAV18717 cDNA encoding GAGE
37	181.2	27.1	538	20	AAV18717 GAGE-2 tumour reje
38	180	26.9	479	22	AAAF68152 Human lung tumour
39	180	26.9	479	24	ABK38063 cDNA encoding clon
40	180	26.9	479	25	ACAL10392 Human lung cancer-
41	180	26.9	479	25	ABX99343 Lung cancer therap
42	179.6	26.8	1245	22	AAAF68152 Human cancer agent
43	178	26.6	461	22	AAAF68152 Human lung tumour
44	178	26.6	461	22	AAAF68152 Human lung tumour
45	178	26.6	461	22	AAAF68152 Human lung tumour
46	178	26.6	461	22	AAAF68152 Human lung tumour
47	169.8	25.4	527	25	ACCS1027 Lung cancer therap
48	169.8	25.4	527	25	ABX76236 Lung cancer-associ
49	169.8	25.4	528	24	ABL66323 Lung cancer relate
50	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
51	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
52	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
53	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
54	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
55	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
56	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
57	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
58	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
59	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
60	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
61	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
62	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
63	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
64	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
65	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
66	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
67	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
68	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
69	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
70	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
71	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
72	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
73	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
74	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
75	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
76	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
77	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
78	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
79	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
80	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE
81	169.8	25.4	532	19	AAV18720 cDNA encoding GAGE

82 123.2 18.4 475 22 AAD14981 Human NOV2 DNA. H
 83 123.2 18.4 538 25 AAT15736 Human cancer/testi
 84 122.4 18.3 124 21 AAC10466 Human secreted pro
 85 122.2 18.3 300 20 AAX55997 Human cDNA clone F
 86 118.8 17.8 215 25 ABZ78639 Tumour suppression
 87 118.8 17.8 215 25 ABZ09186 Human oligonucleot
 88 116.8 17.5 1051 22 AAD14982 Human NOV3 DNA. H
 89 114 17.0 530 22 AAS60104 Human cancer agent
 90 110.4 16.5 365 22 AAS60496 Human cancer agent
 91 110.4 16.5 2168 25 ABZ73737 Secreted protein g
 92 108.8 16.3 2182 21 AAC59112 Human secreted pro
 93 108.8 16.3 2182 25 ABZ73297 Secreted protein-e
 94 108.2 16.2 121724 24 ABQ88143 Human osteoblast d
 95 102.8 15.4 348 22 AAD14997 PCR2.1-83350421-S7
 96 100.8 15.1 201 21 AAD43613 Human secreted exp
 97 99.8 14.9 443 24 ABK64575 Human benign prost
 98 99.2 14.8 346 24 ABK64338 Human benign prost
 99 99.2 14.8 346 24 ABL69505 Prostate cancer re
 100 99.2 14.8 420 21 AAA06691 Human immunogenic

ALIGNMENTS

RESULT 1
 AAK51918
 ID AAK51918 standard; cDNA; 673 BP.

AC AAK51918;
 DT 06-NOV-2001 (first entry)
 DE Human polynucleotide SEQ ID NO 463.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KN tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.

OS Homo sapiens.

XX WO200157190-A2.

PN 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US04098.

XX 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.

PR 20-JUN-2000; 2000US-0598075.

PR 19-JUL-2000; 2000US-0620325.

PR 01-SEP-2000; 2000US-0654936.

PR 15-SEP-2000; 2000US-0683561.

PR 20-OCT-2000; 2000US-0693325.

PR 30-NOV-2000; 2000US-0728422.

XX (HYSE-) HYSEQ INC.

XX WPI; 2001-476283/51.

XX P-PSDB; AAM78785.

XX Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -

XX Claim 1; Page 1690-1681; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the

CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.

XX Sequence 673 BP; 203 A; 145 C; 170 G; 155 T; 0 other;

Query Match 87.0%; Score 581.8; DB 22; Length 673;

Best Local Similarity 95.7%; Pred. No. 2.5e-166;

Matches 617; Conservative 0; Mismatches 7; Indels 21; Gaps 1;

Qy 10 GGCACCGACTTCAGTGTGATGTTCTTGGACACCTGCTCAGTGTGATGTTCTTCTGCGG 59

Db 26 GGCACGAGCTTCAGTGTGATGTTCTTGGACACCTGCTCAGTGTGATGTTCTTCTGCGG 85

Qy 70 CATCTTCCCTTCGACCGCTTTGCGCCAGCTGTTGCGCGCTGGGAGCTGTGAGAGTGTGAG 129

Db 86 CATCTTCCCTTCGACCGCTTTGCGCCAGCTGTTGCGCGCTGGGAGCTGTGAGAGTGTGAG 145

Qy 130 GGGCACGTTCCAGCGCTTCGACTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 199

Db 146 GGGCACGTTCCAGCGCTTCGACTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 205

Qy 190 AGGCGAGTCTCTCCAGGAAGTGAATAGTGAATATAGTGAATATAGTGAATATAGTGAAT 249

Db 206 AGGCGAGTCTCTCCAGGAAGTGAATAGTGAATATAGTGAATATAGTGAATATAGTGAAT 265

Qy 250 TAGCCCTAGCCCAAGAAAGTTTACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 309

Db 266 TAGCCCTAGCCCAAGAAAGTTTACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 320

Qy 310 CTCCCTCTTTGTCGAGGACCCCTGATGATGATGATGATGATGATGATGATGATGATGAT 369

Db 321 -----GACCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 364

Qy 370 AAGTCGGAATCTTACACCTGATCAGAAAGAGAGATGATCAGGCTGCTGCTGCTGCTGCTGCT 429

Db 365 AAGTCGGAATCTTACACCTGATCAGAAAGAGAGATGATCAGGCTGCTGCTGCTGCTGCTGCT 424

Qy 430 AGTCCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGATGATG 489

Db 425 AGTCCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGATGATG 484

Qy 490 TGAAGGTGCTTACTGATGTCAAGGGGAAAGATTCTACCAAAAGCAGAGCACTTTAAATGCC 549

Db 485 TGAAGGTGCTTACTGATGTCAAGGGGAAAGATTCTACCAAAAGCAGAGCACTTTAAATGCC 544

Qy 550 AGAAGCAGGTGAAGGGAATCAAGGTTTAAAGGAAGTAAAGCTGAACACACAAACTG 609

Db 545 AGAAGCAGGTGAAGGGAATCAAGGTTTAAAGGAAGTAAAGCTGAACACACAAACTG 604

Qy 610 TTTTATATATAGATATTTTACTTTTAAAGAGCTCTTAATAAATTTT 654

Db 605 TTTTATATATAGATATTTTACTTTTAAAGAGCTCTTAATAAATTTT 649

RESULT 2
 ABA93876

ID ABA93876 standard; cDNA; 642 BP.

XX AC ABA93876;

XX DT 07-MAY-2002 (first entry)

XX Human G protein-coupled receptor NOV4 encoding cDNA SEQ ID NO:9.

XX XX

XX	CC	The invention comprises a method for diagnosing cancer, the method involves detecting the DNA or protein sequences of human cancer/testis (CT) antigens that are disclosed in the invention. The method of the CC invention is useful for detecting/diagnosing, treating and monitoring a cancer or condition characterised by the expression of a human CT antigen. The present DNA sequence encodes a human CT antigen of the CC invention.
XX	SQ	Sequence 756 BP; 223 A; 150 C; 216 G; 167 T; 0 Other;
		Query Match 83.4%; Score 558; DB 25; Length 756;
		Best Local Similarity 95.8%; Pred. No. 4.5e-159;
		Matches 592; Conservative 0; Mismatches 5; Indels 21; Gaps 1;
QY	37	TGGACACTGCCTCAGGTGTCATGTTCACTCGGCACTTCCCTTCGACCCTTTGCCCAC 96
DB	41	TGGACACTGCCTCAGGTGTCATGTTCACTCGGCACTTCCCTTCGACCCTTTGCCCAC 100
QY	97	GTGTGTACCGCTGGGGAGCTGTGAGAGTGTCAGGGGCAAGTTCACGGCGTCTGACTCTT 156
DB	101	GTGTGTACCGCTGGGGAGCTGTGAGAGTGTCAGGGGCAAGTTCACGGCGTCTGACTCTT 160
QY	157	TCTCTCTCTACTGACCGCAGCCTATAGTCCGCGAGGCCAGTCCCTCCAGGAACCTGAATA 216
DB	161	TCTCTCTCTACTGACCGCAGCCTATAGTCCGCGAGGCCAGTCCCTCCAGGAACCTGAATA 220
QY	217	GTGAATAATGATTTGGCGAGGAAGATCAACATATAGGCTTAGGCCAAGAAGTTTACA 276
DB	221	GTGAATAATGATTTGGCGAGGAAGATCAACATATAGGCTTAGGCCAAGAAGTTTACA 280
QY	277	GCTTCTCTGAGTGTATTGGGGCTATGCTTACTGGCTCCCTTTGTCCCAGGAACCCACATGA 336
DB	281	GCTTCTCTGAGTGTATTGGGGCTATGCTT-----GAAACCACATGA 319
QY	337	TGAAGAGCCTTAAGAAGAGAAACCAACCCACTAAAAGTCGGATCCTACACCTGATFCAGAA 396
DB	320	TGAAGAGCCTTAAGAAGAGAAACCAACCCACTAAAAGTCGGATCCTACACCTGATFCAGAA 379
QY	397	GAGAGAAGATGATCAGGGTGAGCTGAGATTCAGTGCTGACCTGGAGCCGATCTCCA 456
DB	380	GAGAGAAGATGATCAGGGTGAGCTGAGATTCAGTGCTGACCTGGAGCCGATCTCCA 439
QY	457	GGAGCTATGTCAGACAAAGACTGGGGATGGATGTGAAGGTGGTACTGTATGTCAAGGGGAA 516
DB	440	GGAGCTATGTCAGACAAAGACTGGGGATGGATGTGAAGGTGGTACTGTATGTCAAGGGGAA 499
QY	517	GATTTCTACCAAAGCAGAGCACTTTTAAATGCCAGHAGCAGGTGAAGGGAAATCAAGGT 576
DB	500	GATTTCTACCAAAGCAGAGCATTTTTAAATGCCAGHAGCAGGTGAAGGGAAATCAAGGT 559
QY	577	TTAAAGGAGATAAGCTGTAACCAACAACTGTTTTTATATTAGATATTTTACTTTTAA 636
DB	560	TTAAAGGAGATAAGCTGTAACCAACAACTGTTTTTATATTAGATATTTTACTTTTAA 619
QY	637	GAGTCTTAATAAATTTTT 654
DB	620	ATATCTTAATAAAGTTTT 637

KW	developmental disorder; cell signalling disorder;
KW	cell proliferative disorder; cancer; tumour; anaemia;
KW	articular sclerosis; asthma; allergy; diabetes mellitus;
KW	menstrual cycle disorder; bacterial infection; ss.
OS	Homo sapiens.
PN	WO200107471-A2.
XX	
PD	01-FEB-2001.
XX	
PF	21-JUL-2000; 2000WO-US19948.
XX	
PR	21-JUL-1999; 99US-0145075.
PR	08-SEP-1999; 99US-0153129.
PR	10-NOV-1999; 99US-0164647.
XX	
PA	(INCY-) INCYTE GENOMICS INC.
XX	
PI	Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O;
PI	Asimzai Y, Yang J, Lu DAM, Baughn MR, Patterson C, Shah P;
XX	
DR	WPI; 2001-112727/12.
DR	P-PSDB; AAB60500.
XX	
XX	Human cell cycle and proliferation proteins and polynucleotides are
PT	used to treat, diagnose and prevent immune, developmental and cell
PT	signaling disorders and cell proliferative disorders including cancer
XX	
PS	Claim 5; Page 200; 205pp; English.
XX	
CC	Sequences AAF59590-AAF59643 represent cDNAs encoding 54 human
CC	cell cycle and proliferation proteins (CCYPR), AAB60453-AAB60506.
CC	CCYPR and agonists of CCYPR are used to treat diseases or conditions
CC	associated with decreased expression of functional CCYPR, while CCYPR
CC	antagonists are used to treat diseases or conditions associated with
CC	overexpression of functional CCYPR. Monoclonal or polyclonal antibodies
CC	to CCYPR may be used in enzyme-linked immunosorbent assays (ELISA) or
CC	radioimmunoassays to detect CCYPR. CCYPR itself may be used to detect
CC	compounds e.g., antibodies, oligonucleotides and proteins (receptors)
CC	that specifically bind to CCYPR, and in drug screening methods to
CC	identify compounds that modulate the activity of CCYPR. CCYPR
CC	nucleotides can be used to generate transgenic animal models of human
CC	disease, and can be used in gene therapy in target cells with genetic
CC	abnormalities with respect to the expression of CCYPR for the
CC	treatment or prevention of a disorder associated with CCYPR.
CC	Diseases which can be diagnosed, treated and prevented using CCYPR
CC	proteins, nucleic acids, agonists or antagonists include immune,
CC	developmental and cell signalling disorders, and cell proliferative
CC	disorders including cancer. Specific examples of these disorders
CC	include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies,
CC	diabetes mellitus, disorders of the menstrual cycle and infections
CC	caused by bacteria.
XX	
SO	Sequence 580 BP; 167 A; 135 C; 160 G; 118 T; 0 other;

```

Db 191 TAGGTCCGAGCCAGTCTCCAGAACTGAAATAGTGAATATGATTGGCAGGAAG 240
Qy 241 ATCAACATATAGGCTAGGCCAAGAAAGTTTACAGCCTCTCAGCTGTGATTTGGGGCTAT 300
Db 241 ATCAACATATAGGCTAGGCCAAGAAAGTTTACAGCCTCTCAGCTGTGATTTGGGGCTAT 300
Qy 301 GCTTACTGGCTCCCTTTGTCCAGGACCCACTGATGAAGAGCCTTAAGAGAGAAACC 360
Db 301 GCTT-----GAACCCACTGATGAAGAGCCTTAAGAGAGAAACC 339
Qy 361 ACCCACTAAAGTCGGAATCTTACACCTGATCAGAAAGAGAGATGATCAGGTTGCAGC 420
Db 340 ACCCACTAAAGTCGGAATCTTACACCTGATCAGAAAGAGAGATGATCAGGTTGCAGC 399
Qy 421 TGAGATTCAAGTGCCTGAGCCGATCCAGGAGTATCCAGGAGTATGTCAGACAAAGACTGG 480
Db 400 TGAGATTCAAGTGCCTGAGCCGATCCAGGAGTATGTCAGACAAAGACTGG 459
Qy 481 CGATGGATGTGAAGTGGTACTGATGTCAGAGGGAAGATTTTACCAAAGCAGAGCACTT 540
Db 460 GGATGGATGTGAAGTGGTACTGATGTCAGAGGGAAGATTTTACCAAAGCAGAGCACTT 519
Qy 541 TAAATGCCAAGACGAGTGAAGGGAATCACAGCTTTAAGGAAGATAAGCTGAACAA 600
Db 520 TAAATGCCAAGACGAGTGAAGGGAATCACAGCTTTAAGGAAGATAAGCTGAACAA 579
Qy 601 C 601
Db 580 C 580

RESULT 5
AADI4983
ID AADI4983 standard; DNA; 611 BP.
XX
AC AAD14983;
XX
DT 01-NOV-2001 (first entry)
DE
DE Human NOV4 DNA.
XX
KW Human; NOVX; G-antigen; GAGE-like protein; interferon;
KW G-protein coupled receptor; GPCR; hepatocyte nuclear factor;
KW mast cell protease; gene therapy; proliferative disorder; cancer;
KW immune disorder; hepatic disorder; cirrhosis; viral infection;
KW hepatitis; neuroolfactory system-related disorder; neurological disorder;
KW Parkinson's disease; infertility; autoimmune disease; arthritis;
KW multiple sclerosis; allergy; wound healing; cytostatic; nootropic;
KW immunosuppressive; neuroprotective; vulnerary; hepatotropic; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..173
FT FT /*tag= a
FT CDS 174..521
FT FT /*tag= b
FT FT /product= "Human NOV4 protein"
FT 3'UTR 522..611
FT FT /*tag= c
XX
XX WO200161009-A2.
XX
XX 23-AUG-2001.
XX
XX 15-FEB-2001; 2001WO-US04828.
XX
XX 15-FEB-2000; 2000US-0182723.
XX 15-FEB-2000; 2000US-0182724.
XX 15-FEB-2000; 2000US-0182733.
XX 22-FEB-2000; 2000US-0183896.
XX 23-FEB-2000; 2000US-0184275.

```

```

PR 23-FEB-2000; 2000US-0184482.
PR 23-FEB-2000; 2000US-0184497.
PR 24-FEB-2000; 2000US-0184744.
PR 13-APR-2000; 2000US-0197083.
PR 10-AUG-2000; 2000US-0224157.
PR 18-SEP-2000; 2000US-0233405.
PR 27-SEP-2000; 2000US-0236060.
PR 02-JAN-2001; 2001US-0259414.
PR 18-JAN-2001; 2001US-0262454.
PR 14-FEB-2001; 2001US-0783429.
XX
XX (CURA-) CURAGEN CORP.
PA Malyankar UM, Tchernev VT, Padigar M, Taupier RJ, Spytek KA;
XX Majumder K, Guo X, Spaderna SK, Boldog FL;
PI
XX WPI; 2001-514775/56.
DR P-PSDB; AAE08583.
XX
XX Isolated novel polypeptides useful for diagnosis of and treating
PT cancer, infertility, autoimmune diseases, arthritis, multiple
PT sclerosis, allergies, wound healing and hepatic disorders -
XX
XX Claim 9; Page 14; 140pp; English.
XX
XX The present sequence is a human NOV4 DNA. The NOVX protein has homology
CC with one of G-antigen (GAGE)-like protein, interferon, G-protein coupled
CC receptor (GPCR), hepatocyte nuclear factor or mast cell protease. The
CC NOVX is useful for treating or preventing a pathology associated with
CC NOVX. It is also useful for determining the presence or amount of NOVX
CC DNA in a sample, for identifying a potential therapeutic agent and in
CC gene therapy. It is also useful for determining the presence of or
CC predisposition to a disease associated with altered levels of NOVX. It is
CC also useful for the diagnosis and treatment of proliferative disorders,
CC e.g., cancer, immune disorders, hepatic disorders, e.g., cirrhosis, viral
CC infections, e.g., hepatitis, neuroolfactory system-related disorders,
CC neurological disorders, e.g., Parkinson's disease, infertility,
CC autoimmune diseases, arthritis, multiple sclerosis, allergies and wound
CC healing.
XX
XX Sequence 611 BP; 196 A; 127 C; 152 G; 135 T; 1 other;
SQ

```

```

Query Match 75.9%; Score 507.8; DB 22; Length 611;
Best Local Similarity 94.6%; Pred. No. 7.3e-144;
Matches 566; Conservative 0; Mismatches 8; Indels 24; Gaps 3;
Qy 60 GTTCACGTGGGCACTTCCCTTCGACCCCTTTCGCCACGTGGTACCGCT-GGGGAGCTGT 118
Db 9 GTACACTGGGCATCTTCCCTTCGACCCCTTTCGCCACGTGGTACCGCTGGGGAGCTGT 68
Qy 119 GAGAGTGTGAGGGGACGTTCCAGCCGTCTGGACTCTTTCTCTCTACTGAGACGAGCC 178
Db 69 GAGAGTGTGAGGGACAGTTCCAGCCGTCTGGACTCTTTCTCTCTACTGAGACGAGCC 128
Qy 179 TATAGTCCCGAGCCAGTCTCTCCAGGAACCTGAATAGTGAATATGATTCGCGAGGA 238
Db 129 TATAGTCCCGAGCCAGTCTCTCCAGGAACCTGAATAGTGAATATGATTCGCGAGGA 188
Qy 239 AGATCAACATATAGGCTAGGCCCAAGAAAGTTTACAGCCTCTCAGCTGTATTTGGGCT 298
Db 189 AGATCAACATATAGGCTAGGCCCAAGAAAGTTTACAGCCTCTCAGCTGTATTTGGGCT 248
Qy 299 ATGCTTACTGGCTCCCTTTGTCCAGGAACCCCACTGATGAAGAGCTTAAAGAGAGAA 358
Db 249 ATGCTT-----GAACCCACTGATGAAGAGCTTAAAGAGAGAA 287
Qy 359 CCACCCACTAAAGTCGGAATCTTACACCTGA--TCAGAGAGAGAGAGATGATCAGGGTG 416
Db 288 CCACCCACTAAAGTCGGAATCTTACACCTGATCTNAGAGAGAGAGATGATCAGGGTG 347
Qy 417 CAGCTGAGATTCAAGTGCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTGACAGCAAGA 476
Db 348 CAGCTGAGATTCAAGTGCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTGACAGCAAGA 407

```

QY 477 CTGGGATGATGTGAAGGTGGTACTGATGTCAAGGGGAAGATTCTACCAAAAGCAGAGC 536
 Db 408 CTGGGATGATGTGAAGGTGGTACTGATGTCAAGGGGAAGATTCTACCAAAAGCAGAGC 467
 QY 537 ACTTTAAATGCCAAGACGAGGTGAAGGGAATCAAGGTTTAAAGGAAGATAGCTGAA 596
 Db 468 ACTTTAAATGCCAAGACGAGGTGAAGGGAATCAAGGTTTAAAGGAAGATAGCTGAA 527
 QY 597 ACAACAAACTGTTTTTATATTAGATATTTTAAAGAGTCTTAATAAATTTT 654
 Db 528 ACAACAAACTGTTTTTATATTAGATATTTTAAAGAGTCTTAATAAATTTT 585

RESULT 6

AAK52902/c

ID AAK52902 standard; cDNA; 665 BP.

XX AC AAK52902;

XX DT 06-NOV-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 2431.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.

XX OS Homo sapiens.

XX PN WO200157190-A2.

XX PD 09-AUG-2001.

XX XX 05-FEB-2001; 2001WO-US04098.

XX PR 03-FEB-2000; 2000US-0496914.

XX PR 27-APR-2000; 2000US-0560875.

XX PR 20-JUN-2000; 2000US-0598075.

XX PR 19-JUL-2000; 2000US-0620325.

XX PR 01-SEP-2000; 2000US-0854936.

XX PR 15-SEP-2000; 2000US-0863561.

XX PR 20-OCT-2000; 2000US-0693325.

XX PR 30-NOV-2000; 2000US-0728422.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

XX PI Zhao QH, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZH;

XX PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX DR WPI; 2001-476283/51.

XX DR P-PSDB; AAM79769.

XX PT Nucleic acids encoding polypeptides with cytokine-like activities,

XX PT useful in diagnosis and gene therapy -

XX PS Claim 1; Page 4711; 6221pp; English.

XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.

XX CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666

XX CC (AAM80020) are omitted as the relevant pages from the sequence listing

CC were missing at the time of publication.

SQ Sequence 665 BP; 156 A; 170 C; 139 G; 200 T; 0 other;

Query Match 70.1%; Score 469.2; DB 22; Length 665;

Best Local Similarity 88.1%; Pred. No. 4.1e-132;

Matches 580; Conservative 0; Mismatches 43; Indels 35; Gaps 5;

QY 11 GCACCGACTTCAGTGTGCATGTTCTCTTGGACACCTGCTCAGTGTGATGTTCACTGGGC 70

Db 665 GCACCGACTTCAGTGTGCATGTTCTCTTGGACACCTGCTCAGTGTGATGTTCACTGGGC 606

QY 71 ATCTTCCCTTCGACCCCTTTGCCACGCTGTGACCGCTGGGAGCTGTGAGAGTGTGAGG 130

Db 605 ATCTTCCCTTCGACCCCTTTGCCACGCTGTGACCGCTGGGAGCTGTGAGAGTGTGAGG 546

QY 131 GGCACGTTCCAGCCGTTCTGACCTCTTTCTCTCTACTGACACGAGCTTATAGTCCGCA 190

Db 545 GGCACGTTCCAGCCGTTCTGACCTCTTTCTCTCTACTGACACGAGCTTATAGTCCGCA 486

QY 191 GGCACGTTCCAGCCGTTCTGACCTCTTTCTCTCTACTGACACGAGCTTATAGTCCGCA 250

Db 485 GGCACGTTCCAGCCGTTCTGACCTCTTTCTCTCTACTGACACGAGCTTATAGTCCGCA 426

QY 251 AGGCTTAGGCCCAAGAAAGTTTACAGCTCTCTGAGCTGATGGGGCTATGCTTACTGGC 310

Db 425 AGGCTTAGGCCCAAGAAAGTTTACAGCTCTCTGAGCTGATGGGGCTATGCTTACTGGC 372

QY 311 TCCCTTTGTCCTCAGGAAACCCACTGATGAGAGCCCTTAAAGAGAGAAACCCACTTAA 370

Db 371 TCCCTTTGTCCTCAGGAAACCCACTGATGAGAGCCCTTAAAGAGAGAAACCCACTTAA 327

QY 371 AGTGGGAATCTCACCTGATCAGAGAGAGAGATGATCAGGG- - - - -TGCAGCTGAGA 425

Db 326 AGTGGGAATCTCACCTGATCAGAGAGAGAGATGATCAGGGTTGCAGCTTGAGATT 267

QY 426 TTCAGTGGCTTGA- - - - -CCTGGAGCCGATCTCCAGGAGCTATGT- - - - -CAGACAAAGCTG 479

Db 266 TCAAAGTGGCTTGAACCTGGGAGAGCCGATTTTCCAGGAGAGCTATGTTGTCAGACCAAGACTG 207

QY 480 GGGATGGATGTG- - - - -AAGGTGGTACTGATGTCAAGGGGAAGATTCTACCAAAAGCAGAGC 536

Db 206 GGGATGGTGTGTAAGGTGGTACTGATGTCAAGGGGAAGATTCTACCAAAAGCAGAGC 147

QY 537 ACTTTAAATGCCAAGACGAGTGAAGGAAATCAGAGTTTAAAGGAAGATAGCTGAA 596

Db 146 ACTTTAAATGCCAAGACGAGTGAAGGAAATCAGAGTTTAAAGGAAGATAGCTGAA 87

QY 597 ACAACACAACTGTTTTTATATTAGATATTTTAAAGAGTCTTAATAAATTTT 654

Db 86 ACAACACAACTGTTTTTATATTAGATATTTTAAAGAGTCTTAATAAATTTT 29

RESULT 7

AAI58744

ID AAI58744 standard; cDNA; 659 BP.

XX AC AAI58744;

XX DT 22-OCT-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 947.

XX KW Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US34263.
XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX PR 09-JUL-2000; 2000US-0598042.
XX PR 19-JUL-2000; 2000US-0620312.
XX PR 03-AUG-2000; 2000US-0653450.
XX PR 14-SEP-2000; 2000US-0662191.
XX PR 19-OCT-2000; 2000US-0693036.
XX PR 29-NOV-2000; 2000US-0727344.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX DR WPI; 2001-442253/47.
XX DR P-PSDB; AAM39588.
XX PT Novel nucleic acids and polypeptides, useful for treating disorders
XX PT such as central nervous system injuries -
XX PS Claim 1; SEQ ID NO 947; 10078pp; English.
XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX CC the encoded polypeptides (AAM38642-AAM42213) with nootropic
XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful
XX CC in gene therapy. A composition containing a polypeptide or polynucleotide
XX CC of the invention may be used to treat diseases of the peripheral nervous
XX CC system, such as peripheral nervous injuries, peripheral neuropathy and
XX CC localised neuropathies and central nervous system diseases, such as
XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX CC utilisation of the activities such as: immune system suppression,
XX CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX CC assays for receptor activity, arthritis and inflammation, leukaemias and
XX CC C.N.S disorders.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification.
XX SQ Sequence 659 BP; 197 A; 137 C; 173 G; 152 T; 0 other;
Query Match 47.4%; Score 317; DB 22; Length 659;
Best Local Similarity 77.4%; Pred. No. 7.6e-86;
Matches 442; Conservative 0; Mismatches 85; Indels 44; Gaps 3;
QY 86 CCTTTGCCACGTGTGACCGCTGGGAGCTGTGAGGTGTGAGGGGACGCTCCAGCGC 145
DB 109 CTTTTTCCACGTGGTGAATGCCCTGGAGTGTGAGGGTGTGAGGTGCGGTCTCTGCTG 168
QY 146 TCTGACCTTTCT 205
DB 169 TCTGACCTTTCT 210
QY 206 GAACTGAATAGTGAATAGTGTGGGAGAGATCAATACATATAGAGCCCTAGGCCAAGA 265
DB 211 ----TTGCAGTGAATATGATTTTGGCAGGAGATCAATATAGAGCCCTAGGCCGAGG 266
QY 266 AGAAGTTACAGCT 325
DB 267 AGAAGTTACAGCT 305
QY 326 GAACTGAATAGTGAATAGTGTGGGAGAGATCAATACATATAGAGCCCTAGGCCAAGA 385
DB 306 GAGCCCGGTGTGAGGAGCGCTCAGCAAGAGAGAACCCCACTGAAGTTCGGATCCTGCA 365
QY 386 CTTGATCAGAAGAGAGATGATCAGGGTGCAGCTGAGATTCAGTTCAGTTCAGTTCAGT 445

DB 366 CCTGTCAGGAGAGAGAAGATCAGGCTGAGCTCAAGTGCTGACCTGGAA 425
QY 446 GCCGATCTCCAGAGCTATGTCTCAGACAAAGACTGGGATGTGATGAAAGTGTACTGAT 505
DB 426 GCTGATCTCCAGAGCTGTCTCAGTCAAAGACTGGGGGTGAATGTGGAATGTCTCTGAT 485
QY 506 GTCAAGGGGAAGATTCTTACAAAAGCAGAGCACTTTTAAATGCCAGAAGCAGGTGAAGG 565
DB 486 GACACAGGGAGAGATTCTGCCAAATCAGAACATTTAAATGCCAGAAGGAGGTGACAG 545
QY 566 AATCAGAGCTTTAAAGGAGATAGCTGAAACACAC-AAAAGTGTGTATATATAGATA 624
DB 546 CAACCCACAGCTTTAAATGAAGACAGCTGAAACACACAAAAGTGTGTATATATAGATA 605
QY 625 TTTTACTTTAAAGAGCTTTAAATAAATTTTGG 655
DB 606 TTTGACTTTAAATAATCGAAATAAATTTTGG 636
RESULT 8
AAI60530
ID AAI60530 standard; cDNA; 532 BP.
XX AC AAI60530;
XX DT 22-OCT-2001 (first entry)
XX DE Human polynucleotide SEQ ID NO 4519.
XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX KW peripheral nervous system; neuropathy; central nervous system; CNS;
XX KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX KW leukaemia; ss.
XX OS Homo sapiens.
XX PN WO200153312-A1.
XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US34263.
XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX PR 09-JUL-2000; 2000US-0598042.
XX PR 19-JUL-2000; 2000US-0620312.
XX PR 03-AUG-2000; 2000US-0653450.
XX PR 14-SEP-2000; 2000US-0662191.
XX PR 19-OCT-2000; 2000US-0693036.
XX PR 29-NOV-2000; 2000US-0727344.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX DR WPI; 2001-442253/47.
XX DR P-PSDB; AAM41374.
XX PT Novel nucleic acids and polypeptides, useful for treating disorders
XX PT such as central nervous system injuries -
XX PS Claim 1; SEQ ID NO 4519; 10078pp; English.
XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful
XX CC in gene therapy. A composition containing a polypeptide or polynucleotide
XX CC of the invention may be used to treat diseases of the peripheral nervous
XX CC system, such as peripheral nervous injuries, peripheral neuropathy and
XX CC localised neuropathies and central nervous system diseases, such as
XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX CC utilisation of the activities such as: immune system suppression,
XX CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX CC assays for receptor activity, arthritis and inflammation, leukaemias and
XX CC C.N.S disorders.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification.

RESULT 9	
ABX77605	
ID	ABX77605 standard; CDNA; 509 BP.
XX	
AC	ABX77605;
XX	
DT	09-APR-2003 (first entry)
XX	
DE	Differentially expressed breast cancer associated cDNA #100.
XX	
KW	Breast cancer; differential gene expression; BC-cDNA;
KW	breast cancer diagnosis; breast cancer monitoring;
KW	breast cancer treatment; breast cancer staging; gene; ss.
XX	
OS	Unidentified.

XX	Sequence	509 BP; 158 A; 100 C; 139 G; 112 T; 0 other;
	Query Match	43.4%; Score 290.2; DB 25; Length 509;
	Best Local Similarity	76.9%; Pred. No. 9.3e-78;
	Matches 419; Conservative	0; Mismatches 68; Indels 58; Gaps 3;
QY	112	GAGCTGTCAGAGTGTGAGGGGCACGCTCCAGCGCTGGAGCTCTTTCTCTCTCTACTGTAGA 171
DB	1	GAGTTGTGAGGGTGTGAGGTCGCGTCTCTGCTGTCTGGACTTTTCTGCCACTGTAGA 60
QY	172	CGCAGCCTATAGGTCGCGAGCGCAGTCCTCCAGGAATGGAANTGTGAATATGAGTTG 231
DB	61	CGCAGCTGT-----GTGAATATGATGTTG 84
QY	232	CGGAGGAGATCAACATATAGCCCTAGCCRAGAAGAGTTTACAGCCCTCTGAGCTGAT 291
DB	85	GGGAGAAGATCAACATATAGCCCTAGGCCGAGGAGAGTGTACACCTCTTGAGCTGAT 144
QY	292	TGGGGCTATGCTTACTTGGCTCCCTTTGTGCCAGGAACCCACTGATGAAGAGCCCTAAAGA 351
DB	145	TGGGCTATGCT-----GGAGCCGGTGTAGGAGCCTCAGCA 183
QY	352	AGAGAAACCCACCACTAAAGTCGGATCCCTACCTGATCAGAGAGAGAGATGATCA 411
DB	184	AGAGAAACCCACCACTAAAGTCGGATCCCTGACCTGGTCAGGAGAGAGAGAGATCA 243
QY	412	GGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTCAGAC 471
DB	244	GGGTGCAGCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCAGTC 303
QY	472	AAAGACTGGGGATGGATGTGAGGTGGTACTGATGTCAAGGGGAGATTCTACCAAAGC 531
DB	304	AAAGACTGGGGGTGAATGTGGAAATGGTCTGATGCCAGGGGAGAGATTCTGCCAAATC 363
QY	532	AGAGCACTTTAAATATGCCAAGACGAGGTGAAGGAAATCACAGGTTTAAAGGAAGATAAG 591

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (II) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (III). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 503 BP; 158 A; 98 C; 121 G; 105 T; 21 other;

Query Match 42.0%; Score 280.8; DB 23; Length 503;
Best Local Similarity 81.1%; Pred. No. 6.7e-75;
Matches 360; Conservative 0; Mismatches 62; Indels 22; Gaps 2;

QY 213 AATAGTGAATATGAGTGGCGAGGAAGATCAACATATAGGCTAGGCCAAGAAGATT 272
DB 49 AATGTGAATATGATTTGGCGAGGAAGATCAACATATAGGCTAGGCCGAGGAAGTG 108

QY 273 TACAGCTCTGAGCTGATGGGCTATGCTTACTGGTCCCTTTGTCGCCAGAACCA 332
DB 109 TACCACCTCTGAGCTGATTTGGGCTATGCT-----GGAGCCCG 147

QY 333 CTGATCAAGAGCTCAAGAGAGAGAACCCACCTAAAGTCGGATCTACACCTGATC 392
DB 148 GTGATGAGAGCTCTCAGAGAGAGAACCCACCTGAAAGTCGGATCTGCACTGGTC 207

QY 393 AGAAGAGAGAGATGATCAGGCTGAGATTCAGATTCAGTGCCTGACCTGGAGCGGATC 452
DB 208 AGGAGAGAGAGAGATCAGGCTGAGATTCAGATTCAGTGCCTGACCTGGAGCGGATC 267

QY 453 TCCAGAGGCTATCTCAGACAAAGACTGGGATGATGATGATGATGATGATGATGATG 512
DB 268 TCCAGAGGCTATCTCAGTCAAGACTGGGATGATGATGATGATGATGATGATGATG 327

QY 513 GGAAGATTCTACCAAGAGAGACACTTTAAATGCCAGAGCAGGTGAAGGGAATCAC 572
DB 328 GGAAGATTCTGCCAAATATCAGAACAAATTTAAATGCCAGAGGAGGTGACAGCAACCAC 387

QY 573 AGTTTAAAGGAAGATGAGTGAACCAACAC-AAAAGCTTTTATATTAGATATTTTACT 631
DB 388 AGTTTAAATGAGCAGCTGAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 447

QY 632 TTAAGAGCTCTTAATAAATTTTG 655
DB 448 TAAATATATCAATAAATTTTG 471

RESULT 12
AAS69486
ID AAS69486 standard; cDNA; 661 BP.
XX
AC AAS69486;
XX
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #5290.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.
XX OS
XX PN WO200175067-A2.
XX
XX PD 11-OCT-2001.
XX
XX PF 30-MAR-2001; 2001WO-US08631.
XX
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX PI Drmanac RT, Liu C, Tang YT;
XX P-PSDB; ABG05299.
XX WPI; 2001-639362/73.
XX
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity
XX
XX Claim 1; SEQ ID No 5290; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (III). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 661 BP; 212 A; 129 C; 158 G; 162 T; 0 other;

Query Match 36.2%; Score 242.4; DB 23; Length 661;
Best Local Similarity 81.5%; Pred. No. 3.6e-63;
Matches 362; Conservative 0; Mismatches 56; Indels 26; Gaps 5;

QY 217 GTGAATATGATGTTGGCGAGGAAGATCAACATATAGGCTAGGCCAAGAAGATTTACA 276
DB 222 GTGAATATGATGTTGGCGAGGAAGATCAACATATAGGCTAGGCCAAGAAGATTGACC 281

QY 277 GCCTCTCAGCTGATTTGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAACCCACTGA 336
DB 282 ACCTCTCAGCTGATTTGGGCTATGCT-----GGAGCCCGGTGA 320

QY 337 TGAAGAGCTTAAGAGAGAGAACCCACCTAAAGTCGGATCTTACACCT-GATCAGA 395
DB 321 TGAGAGGCTCAGCAAGAGAGAACCCACCTAAAGTCGGATCTTACACCTGGGTGAGG 380

QY 396 AGAGAGAAGATGATCAGGCTCAGCTGAGATTCAGTGCCTGACCTGAGAA-GCCGATCTC 454
DB 381 AGAGAGAAGAGATCAGGCTCAGCTGAGATTCAGTGCCTGACCTGAGAGGCTGATCTC 440

QY 455 CAGGAGCTATGTCAGACAAAGACTGGGATGATGATGATGATGATGATGATGATGATG 514
DB 441 CAGGAGCTCTCTCAGTCAAGACTGGGATGATGATGATGATGATGATGATGATGATG 500

QY 515 AAGATTCTACCAAGCAGAGC-ACCTTAAATGCCAGAA-GCAGGTGAAGGGAAATCAC 572
D501 AAGATTCTGCCAAATCAGAACAATTTTAAATGCCAGAGGGAGGTGACAGGCAACCAC 560
QY 573 AGGTTTAAAGGAGATAGCTGAACACAC-AAAAGTGTCTTTTATATTAGATATTTTACT 631
D561 AGGTTTAAATGAAGCAGCTGAACACACAAACAACTGTTTATCTAAGATATTTGACT 620
QY 632 TTAAGAGTCTTAATAATTTTGTG 655
D621 TAAATAATATCGAATAAACTTTTG 644

RESULT 13
AAH83260
ID AAH83260 standard; cDNA; 320 BP.
XX
AC AAH83260;
XX
DT 25-SEP-2001 (first entry)
XX
DE Human ovarian tumour associated polynucleotide sequence SEQ ID NO:884.
XX
KW Human; ovarian tumour; ovarian cancer; diagnosis; gene therapy;
KW immunogenic; vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO200151513-A2.
XX
PD 19-JUL-2001.
XX
PF 16-JAN-2001; 2001WO-US01575.
XX
PR 14-JAN-2000; 2000US-0176722.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA;
XX
DR WPI; 2001-425866/45.
XX

Novel ovarian tumor proteins, and nucleic acids encoding them, used to treat and diagnose cancers, particularly ovarian cancer -
Claim 5; Page 225; 338pp; English.

AAH82377 to AAH83878 represent human ovarian tumour-associated polynucleotide sequences which encode ovarian tumour proteins. The ovarian tumour protein and polynucleotide sequences have cytostatic activity, and can be used in gene therapy and vaccine production. The ovarian tumour proteins and polynucleotides can be used to inhibit the development of cancer, particularly ovarian cancer. They can also be used to diagnose the onset and progression of cancer.

Sequence 320 BP; 96 A; 65 C; 85 G; 74 T; 0 other;
Query Match 32.1%; Score 215; DB 22; Length 320;
Best Local Similarity 100.0%; Pred. No. 5.3e-55;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 420 CTGAGATTCAAGTGGCTGACCTGGAGCGGATCTCCAGGAGCTATGTCCAGCAAAAGACTG 479
D1 CTGAGATTCAAGTGGCTGACCTGGAGCGGATCTCCAGGAGCTATGTCCAGCAAAAGACTG 60
QY 480 GGGATGGATGTAAGTGGTACTGATGTCAGGGGAGATCTACCAAAAGCAGGCACT 539
D61 GGGATGGATGTAAGTGGTACTGATGTCAGGGGAGATCTACCAAAAGCAGGCACT 120
QY 540 TTAATAATGCCAAGCAGGTGAAGGGAATCAGAGTTTAAAGGAAGATAAGCTGAACA 599
D121 TTAATAATGCCAAGCAGGTGAAGGGAATCAGAGTTTAAAGGAAGATAAGCTGAACA 180

QY 600 ACACAACTGTTTATATATAGATATTTTACTTTA 634
D181 ACACAACTGTTTATATATAGATATTTTACTTTA 215

RESULT 14
AAS24637
ID AAS24637 standard; cDNA; 214 BP.
XX
AC AAS24637;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human ovarian PCR-subtracted cDNA library clone #818.
XX
KW Immunogenic protein; cancer; ovarian tumour; T-cell stimulation; ss;
KW gene therapy; cytostatic; T-cell expansion; nucleic acid hybridisation;
KW primer; probe.
XX
OS Homo sapiens.
XX
PN WO200157207-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US03733.
XX
PR 04-FEB-2000; 2000US-0180403.
XX
PR 28-MAR-2000; 2000US-0192745.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Mannion J;
XX
DR WPI; 2001-488879/53.
XX
PT New polynucleotides encoding ovarian tumour proteins, useful for treating ovarian cancer, and as probes, primers, and markers of cancer progression -
PT
PS Example 1; page 241-242; 378pp; English.

The invention comprises compositions used for the therapy and diagnosis of ovarian cancer. The compositions comprise one or more ovarian tumour proteins, their associated polynucleotides, or immunogenic portions of the proteins. The ovarian tumour polynucleotides and polypeptides are useful for stimulating and/or expanding T cells specific for a tumour protein. They are also useful for inhibiting the development of cancer in a patient with an ovarian tumour DNA or protein by incubating isolated T-cells allowing them to proliferate, and administering to the patient. The sequences can be used as markers for cancer, for example, to monitor ovarian cancer progression. Probes and primers are useful in nucleic acid hybridisation, in detecting the presence of complementary sequences in a given sample, for preparing mutant species and for preparing other genetic constructions. Sequences AAS23820-AAS25231 and AAS25238-AAS25549 represent human ovarian tumour protein cDNA clones.

Sequence 214 BP; 74 A; 34 C; 55 G; 51 T; 0 other;
Query Match 32.0%; Score 214; DB 22; Length 214;
Best Local Similarity 100.0%; Pred. No. 8.7e-55;
Matches 214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 420 CTGAGATTCAAGTGGCTGACCTGGAGCGGATCTCCAGGAGCTATGTCCAGCAAAAGACTG 479
D1 CTGAGATTCAAGTGGCTGACCTGGAGCGGATCTCCAGGAGCTATGTCCAGCAAAAGACTG 60
QY 480 GGGATGGATGTAAGTGGTACTGATGTCAGGGGAGATCTACCAAAAGCAGGCACT 539
D61 GGGATGGATGTAAGTGGTACTGATGTCAGGGGAGATCTACCAAAAGCAGGCACT 120
QY 540 TTAATAATGCCAAGCAGGTGAAGGGAATCAGAGTTTAAAGGAAGATAAGCTGAACA 599

Db 121 TTAATAATGCCAAGCAGGTGAAGGAAATCAACAGGTTTAAAGGAAGATAAGCTCAACA 180
 QY 600 ACACAACTGTTTATATAGATATTTACTTT 633
 Db 181 ACACAACTGTTTATATAGATATTTACTTT 214
 RESULT 15
 ID AAF68151 standard; cdna; 399 BP.
 XX AAF68151;
 AC AAF68151;
 XX 12-APR-2001 (first entry)
 DT Human lung tumour protein related nucleotide sequence SEQ ID NO:69.
 DE Human; lung cancer; lung tumour; lung tumour protein; gene therapy;
 KW lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;
 KW cytostatic; antisense inhibition; ss.
 XX Homo sapiens.
 OS
 XX WO200100828-A2.
 PN 04-JAN-2001.
 XX 30-JUN-2000; 2000WO-US18061.
 XX 30-JUN-1999; 99US-0346492.
 PR 15-OCT-1999; 99US-0419356.
 PR 17-DEC-1999; 99US-0466867.
 PR 30-DEC-1999; 99US-0476300.
 PR 06-MAR-2000; 2000US-0519642.
 PR 22-MAR-2000; 2000US-0533077.
 PR 10-APR-2000; 2000US-0546259.
 PR 27-APR-2000; 2000US-0560406.
 PR 05-JUN-2000; 2000US-0589184.
 XX (CORI-) CORIXA CORP.
 PA Wang T, Bangur CS, Lodes MJ, Fanger GR, Vedvick TS, Carter D;
 FI Retter MW, Mannion J;
 PI WPI; 2001-071486/08.
 DR Lung tumor-associated proteins and the nucleic acids that encode them,
 XX useful for preventing, diagnosing and treating lung cancer -
 FS Claim 4; Page 173; 436pp; English.
 XX The present invention describes immunogenic portions of lung tumour-
 CC associated proteins (I) and the nucleic acids (NAs) that encode them.
 CC (I) have cytostatic activity and can be used in gene therapy, antisense
 CC inhibition and in vaccines. The NAs and the lung tumour-associated
 CC proteins they encode may be used in the prevention, treatment and
 CC diagnosis of diseases associated with their inappropriate expression,
 CC especially lung cancers. For example, the NAs may be administered to
 CC treat diseases by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of the protein by expressing inactive proteins
 CC or to supplement the patients own production of (I). Additionally, the
 CC NAs may be used to produce the lung-tumour associated protein, according
 CC to standard recombinant DNA methodology. Conversely, antisense NA
 CC molecules may be administered to down regulate protein expression by
 CC binding with the cells own genes and preventing their expression. The NA
 CC and complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar NA sequences in
 CC samples, and hence which patients may be in need of treatment for lung
 CC cancer. The (I) may be used as antigens in the production of antibodies
 CC and in assays to identify modulators (agonists and antagonists) of the
 CC expression and activity of the protein. AAF68083 to AAF68878 and
 CC AAB76848 to AAB76878 represent human lung tumour protein related

CC nucleotide and protein sequences which are used in the exemplification
 CC of the present invention.
 XX Sequence 399 BP; 150 A; 87 C; 94 G; 67 T; 1 other;
 QY Query Match 30.5%; Score 204; DB 22; Length 399;
 Best Local Similarity 77.4%; Pred. No. 1.3e-51;
 Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
 QY 319 GTCCAGGAAACCCACTGATGAAGAGCTTAAAGAGAGAAACCCACCTAAAGTGGAA 378
 Db 40 GTCCAGGAGCCAGTAATGGAGAGCCCAAAAGAAAGAACCCAGCAGCTGAAGTCGGGA 99
 QY 379 TCCTACACCTGATCAGAGAGAGAGAGATGATCAGGCTGAGCTGAGATTCAGTGCCTGA 438
 Db 100 TCCTACACCTGGGAGAGAGAGAGATCAGGATCAGCTGAGATTCAGTGCCTGA 159
 QY 439 CCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGATGGGGATGGATGGAAGGT-G 497
 Db 160 CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGG 219
 QY 498 GTACTGATCTCAAGGGGAGAGATTCTACCAAAAGCAGAGCACTTTAAAAATGCCAGAGCAG 557
 Db 220 GTTCCGGCTCAAGGTGAAGATAATACCTTAAAGAGAACACTGTAAATGCCAGAGCAG 279
 QY 558 GTGAAGGGAATCACAGGTTTAAAGGAAGATAAGCTGAAACACACAACTGTTTTTATA 617
 Db 280 GTGAAGAGCAACACCAAGTTTAAATGAAGACAAGCTGAAACACCGCAAGCTGGTTTTATA 339
 QY 618 TTAGATATTTTACTTTTAAAGAGTCTTAATAAATTTT 653
 Db 340 TTAGATATTTTACTTTTAAAGAGTCTTAATAAATTTT 375
 RESULT 16
 ID ABK38062 standard; cdna; 399 BP.
 XX ABK38062;
 AC
 XX 21-MAY-2002 (first entry)
 DT cdna encoding clone #19107 (L552S) of lung tumour protein.
 DE Lung tumour; cancer; T cell; immune response stimulator;
 KW cytostatic; gene; ss.
 XX Homo sapiens.
 OS
 XX WO200204514-A2.
 PN 17-JAN-2002.
 XX 10-JUL-2001; 2001WO-US22059.
 XX 11-JUL-2000; 2000US-0614124.
 PR 29-AUG-2000; 2000US-0651563.
 PR 08-SEP-2000; 2000US-0658824.
 PR 26-SEP-2000; 2000US-0671325.
 PR 06-OCT-2000; 2000US-0677419.
 PR 30-OCT-2000; 2000US-0702705.
 PR 13-DEC-2000; 2000US-0736457.
 PR 03-MAY-2001; 2001US-0849626.
 XX (CORI-) CORIXA CORP.
 PA Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;
 PI Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS;
 PI McNabb A, Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;
 XX WPI; 2002-164634/21.
 DR Novel polynucleotide encoding a lung tumour polypeptide useful for
 XX PT

PA	(FANG/) FANGER G R.	
PA	(WANG/) WANG A.	
PA	(WANG/) WANG T.	
PA	(SWIT/) SWITZER A P.	
PA	(MCNE/) MCNEILL P D.	
PA	(CLAP/) CLAPPER J D.	
XX		
XX	Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;	
PI	Clapper JD;	
XX		
XX	WPI; 2003-352750/33.	
XX		
XX	Novel lung cancer polynucleotide encoding lung cancer protein, useful	
PT	for detecting the presence of lung cancer in a patient, and in	
PT	pharmaceutical compositions, e.g. vaccines, for treating lung cancer	
PT		
XX		
PS	Example 1; Page -: 72pp; English.	
XX		
XX	The invention relates to a polynucleotide encoding a lung tumour protein,	
CC	comprising a sequence selected from any of the 14 sequences	
CC	mentioned in the specification, or a sequence (S2) mentioned in	
CC	specification, complement of S1, sequences consisting of at least 20	
CC	contiguous residues of S1, sequences that hybridise to S1, sequences	
CC	having 75% preferably 90%, identity to S1, or degenerate variants of	
CC	S1. Also included are an isolated polypeptide (comprising a sequence (S3)	
CC	selected from any one of the 4 amino acid sequences mentioned in the	
CC	specification, a sequence encoded by the polynucleotide, or sequences	
CC	having at least 70%, preferably 90%, identity to a sequence encoded by	
CC	the polynucleotide), an expression vector comprising the polynucleotide	
CC	operably linked to an expression control sequence, a host cell	
CC	transformed or transfected with the vector, an isolated antibody (or its	
CC	antigen-binding fragment) that specifically binds to the polypeptide,	
CC	detecting the presence of a cancer in a patient, a fusion protein	
CC	comprising the polypeptide, an oligonucleotide that hybridises to	
CC	S1 under moderately stringent conditions, stimulating and/or expanding T	
CC	cells specific for a tumour protein (comprising contacting T cells with	
CC	the polynucleotide, protein or antigen-presenting cells, under conditions	
CC	and for a time sufficient to permit the stimulation and/or expansion of T	
CC	cells) and inhibiting the development of a cancer in a patient (by	
CC	incubating CD4 ⁺ and/or CD8 ⁺ T cells isolated from a patient with the	
CC	polynucleotide, protein or antigen presenting cells that express the	
CC	polynucleotide, such that T cells proliferate, administering to the	
CC	patient an effective amount of the proliferated T cells, and thus	
CC	inhibiting the development of a cancer in the patient. The	
CC	polynucleotide, protein and cells are useful in a composition for	
CC	stimulating an immune response in a patient, and for treating a cancer in	
CC	a patient (particularly lung cancer). The oligonucleotide is useful for	
CC	determining the presence of a cancer in a patient. The protein and	
CC	oligonucleotides are useful in pharmaceutical compositions, e.g.	
CC	vaccines. The polynucleotide is also useful as a probe or primer for	
CC	nucleic acid hybridisation, and in the design and preparation of	
CC	ribosome molecules for inhibiting expression of tumour polypeptides and	
CC	proteins in tumour cells. An amplified portion of the polynucleotide is	
CC	useful for isolating a full-length gene from a suitable library.	
CC	The present sequence is a cDNA (full length, extended or partial)	
CC	isolated from a library derived from lung tumour/cancer cells.	
CC	Note: The sequence data for this patent did not form part	
CC	of the printed specification, but was obtained in electronic	
CC	format directly from the USPTO	
CC	at seqdata.uspto.gov/sequence.html?DocId=20020197669.	
XX		
XX	Sequence 399 BP; 150 A; 87 C; 94 G; 67 T; 1 other;	
SQ		
	Query Match 30.5%; Score 204; DB 25; Length 399;	
	Best Local Similarity 77.4%; Pred. No. 1.3e-51;	
	Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;	
QY	319 GTCCAGGACCCACTGATGACAGAGCCTAAGAGAGAGAACCCACCCTAAAGTCGGAA 378	
Db	40 GTCCCCAGGACCCAGTAATGGAGAGCCCAAAAAGAGAACCCAGCAGCTGAAGAGTCGGGA 99	
QY	379 TCCTACACCTGTATCAGAGAGAGAGATGATCATCGGGTGCAGCTTCAGATTCAAGTGCCTGA 438	

100	TCCTACACCTGGGCAGCAGACAGAAAGATCAGGTACAGCTGAGATCCAGCTGGCGCA	159
439	CCTGGAACCCGACTCTCCAGGAGCTATGTCAGACAAGACATGGGATCGGATGTGAAGGT-G	497
160	CATGGAAGGTGATCTTGCAAGAGCTGCATCAGTCAACACCGGGATAAATCTGGATTTGG	219
438	GTACTGATGTCAAGGGGAAGATTCTACCAAAAGCAGAGCACTTTAAAAATGCCAGAACGAG	557
220	GTTCGGCGGTCAAGGTGAAGATAATACCTAAAGAGGAAACACTGTAAATATGCCAGAACGAG	279
558	GTGAAGGGAATTCACAGGTTTTAAAGGAAGATAAGCTGAAAACAACAACACTGTTTTTATA	617
280	GTGAAGACCAACCAACAGTTTTTAATGAAGACAAGCTGAACACACGACGCTGTTTTATA	339
618	TTAGATATTTCATTTAAAGAGTCTTTAATAAATTTT	653
340	TTAGATATTTCAGCTTAAACTATCTCAATAAAGTTTT	375

RESULT 18	
ABX99342	
ID	ABX99342 standard; cDNA; 399 BP.
XX	
XX	ABX99342;
XX	
XX	22-MAY-2003 (first entry)
XX	
XX	Lung cancer therapyand diagnosis associated cDNA #69.
DE	
XX	
XX	Lung cancer; cytostatic; vaccine; gene therapy; cancer;
KW	
KW	gene; ss.
XX	
XX	
OS	Homo sapiens.
XX	
PN	US2002172952-A1.
XX	
XX	
PD	21-NOV-2002.
XX	
XX	
PF	10-JUL-2001; 2001US-0902941.

PR	30-JUN-1999;	98US-0346492.
PR	15-OCT-1999;	99US-0419356.
PR	17-DEC-1999;	99US-0466867.
PR	30-DEC-1999;	99US-0476300.
PR	06-MAR-2000;	2000US-0519642.
PR	22-MAR-2000;	2000US-0533077.
PR	10-APR-2000;	2000US-0548259.
PR	27-APR-2000;	2000US-0560406.
PR	05-JUN-2000;	2000US-0589184.
PR	11-JUL-2000;	2000US-0614124.
PR	29-AUG-2000;	2000US-0651563.
PR	08-SEP-2000;	2000US-0658824.
PR	26-SEP-2000;	2000US-0671325.
PR	06-OCT-2000;	2000US-0677419.
PR	30-OCT-2000;	2000US-0702705.
PR	13-DEC-2000;	2000US-0736457.
PR	03-MAY-2001;	2001US-0849626.
XX		
XX	(CORI-) CORIXA CORP.	
PA	Henderson RA, Wang T, Watanabe Y,	
PA	Carter D, Fanger GR,	
PI	Johnson JC, Retter MM,	Durham M;
PI	Bangor CS, Menabba A;	
XX		
XX	WPI; 2003-328427/31.	
DR		

XX New polynucleotide, useful for preparing a composition for treating or
PT inhibiting development of cancer, e.g. lung cancer -
PT
XX Example 1; SEQ ID NO 69; 82pp; English.
XX
XX The invention describes an isolated polynucleotide comprising one of 32
CC sequences, complement or degenerate variants of them. The polynucleotide
CC is useful for preparing a composition e.g. a vaccine or for gene therapy.
CC

	for treating or inhibiting development of cancer, e.g. lung cancer.
CC	This sequence represents a polynucleotide associated with the
CC	compositions and methods for the therapy and diagnosis of lung cancer.
XX	
Sequence 399 BP; 150 A; 87 C; 94 G; 67 T; 1 other;	
Query March	30.5%; Score: 204; DB 25; Length 399;
Best Local Similarity	77.4%; Pred. No. 1.3e-51;
Matches 260; Conservative	0; Mismatches 75; Indels 1; Gaps 1;
319 GTCCCAGGAACCCACTGATGAAGAGCCTAAAGAAGAGAAACCACCCCACTAAAAGTCGGAA	378
40 GTCCCAGGAGCCCCAGTAAATGGAGAGCCCCAATAAAGAACCCAGCAGCTGAAGATCGGGA	99
379 TCCTACACCTGATCAGAAGAGAGAGATGATCAGGGTGACGTGAGATTCAAGTGCCTGA	438
100 TCCTACACCTGGCGCACAGACAAGAAGATCAGGATACGCTAGATGCCAGTGCSCGA	159
439 CCTGGAAGCCGATCTCCAGGAGCTATGTGAGACAAGACTGGGATGGATGTGAAGT-G	497
160 CATGGAAGGTGATCTCCAAGAGCTCATCAGTCAACACCGGGGTAATCTCGATTGG	219
498 GTACTGATGTCAAGGGGAAGATTCTACMAAAGCAGAGACACTTTTAAATGCCAGAAGCAG	557
220 GTTCCGCGCTCAAGGTTGAAGTAATACTCTAAAGAGGAACTGTGTAATATGCCAGAAGCAG	279
558 GTGAAGGGAATCACAAGGTTTTAAAGGAAGATAAGCTGAAACCAACACAAACTGTTTTATA	617
280 GTGAGAGCAACCAACAGTTTTAAATGAAGACAGCTGAAAACAACCGACGCTGTTTTATA	339
618 TTGATATNTTACTTTAAAGAGCTCTTAATAAATTTT	653
340 TTGATATNTTACTTTAAACTATCTCAATAAAGTTTTT	375

RESULT 19	
AAF68851	
IID	AAF68851 standard; cDNA; 457 BP.
XX	
XX	AAF68851;
XX	
XX	
DT	12-APR-2001 (first entry)
XX	
XX	Human lung tumour protein related nucleotide sequence SEQ ID NO:790.
XX	
DE	Human; lung cancer; lung tumour; lung tumour protein; gene therapy;
KW	lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;
KW	cytostatic; antisense inhibition; ss.
KW	
XX	
XX	Homo sapiens.
OS	
XX	
XX	WO200100828-A2.
XX	
XX	04-JAN-2001.
XX	
XX	30-JUN-2000; 200WO-US18061.
XX	
XX	30-JUN-1999; 99US-0346492.
PR	
PR	15-OCT-1999; 99US-0419356.
PR	17-DEC-1999; 99US-0466867.
PR	30-DEC-1999; 99US-0476300.
PR	06-MAR-2000; 2000US-0519642.
PR	22-MAR-2000; 2000US-0533077.
PR	10-APR-2000; 2000US-0546259.
PR	27-APR-2000; 2000US-0560406.
PR	05-JUN-2000; 2000US-0589184.
XX	
XX	(CORI-) CORIXA CORP.
PA	
XX	
XX	Wang T, Bangur CS, Lodes MJ, Fanger GR, Vedvick TS, Carter D;
PI	Rettner MW, Mannion J;
PI	
XX	
XX	WPI: 2001-071498/08.
DR	

XX Lung tumor-associated proteins and the nucleic acids that encode them,
PT useful for preventing, diagnosing and treating lung cancer -
XX
XX
XX Claim 4; Page 412; 436pp; English.
XX
XX The present invention describes immunogenic portions of lung tumour-
CC associated proteins (I) and the nucleic acids (NAs) that encode them.
CC (I) have cytostatic activity and can be used in gene therapy, antisense
CC inhibition and in vaccines. The NAs and the lung tumour-associated
CC proteins they encode may be used in the prevention, treatment and
CC diagnosis of diseases associated with their inappropriate expression,
CC especially lung cancers. For example, the NAs may be administered to
CC treat diseases by rectifying mutations or deletions in a patient's genome
CC that affect the activity of the protein by expressing inactive proteins
CC or to supplement the patient's own production of (I). Additionally, the
CC NAs may be used to produce the lung-tumour associated protein, according
CC to standard recombinant DNA methodology. Conversely, antisense NA
CC molecules may be administered to down regulate protein expression by
CC binding with the cells own genes and preventing their expression. The NA
CC and complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar NA sequences in
CC samples, and hence which patients may be in need of treatment for lung
CC cancer. The (I) may be used as antigens in the production of antibodies
CC and in assays to identify modulators (agonists and antagonists) of the
CC expression and activity of the protein. AAF68083 to AAF68878 and
CC AAB76848 to AAB76878 represent human lung tumour protein related
CC nucleotide and protein sequences which are used in the exemplification
CC of the present invention.
XX
XX Sequence 457 BP; 146 A; 105 C; 121 G; 85 T; 0 other;
SQ

Query Match 30.5%; Score 204; DB 22; Length 457;
Best Local Similarity 77.4%; Pred. No. 1.4e-51;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
XX 319 GTCCAGGAAACCATGTGATGAAGAGCCCTTAAGAGAGAAACACCCATTAAGTCGGAA 378
DB 120 GTCCAGGAGCCCGAGTAATGGAGAGCCGCCAAAAGAGAGACGAGCTGAAGTCGGGA 179
XX 379 TCTACACCTGATCAGAGAGAGAGATGATCAGGCTGCAGTGCAGTTCAGTTCAGTTCGA 438
DB 180 TCTACACCTGGCAGCAGACAGAGAGATCAGATCAGTGCAGTTCAGTTCAGTTCGA 239
XX 439 CTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGATGAGTGAAGGT-G 497
DB 240 CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGG 299
XX 498 GTACTGATCTCAAGGGGAGATTTCTACCAAAAGCAGAGACACTTTAAATGCCAGAGCAG 557
DB 300 GTTCCGGCTCAAGGTGAAGATAATACCTTAAGAGAGAACTGTGTAATGCCAGAGCAG 359
XX 558 GTGAAGGGAATCAGAGGTTTAAAGGAAGATAGCTGAAACCAACACAACTGTTTTATA 617
DB 360 GTGAAGAGCAACCAAGTTTAAATGAAGCAAGCTGAAACCAACGCAAGCTGGTTTTATA 419
XX 618 TTAGATATTTTACTTTAAGAGTCTTAATAATTTT 653
DB 420 TTAGATATTTGACTTAACTATCTCAATAAAGTTTT 455

RESULT 20
ABK38762
ID ABK38762 standard; cDNA; 457 BP.
XX
XX ABK38762;
AC
XX
XX 21-MAY-2002 (first entry)
DT
XX cDNA encoding L552S lung tumour protein.
DE
XX Lung tumour; cancer; T cell; immune response stimulator;
XX cytosstatic; gene; ss.
XX

XX Homo sapiens.
XX WO200204514-A2.
XX
XX 17-JAN-2002.
XX
XX 10-JUL-2001; 2001WO-US22058.
XX
XX 11-JUL-2000; 2000US-0614124.
XX 29-AUG-2000; 2000US-0651563.
XX 08-SEP-2000; 2000US-0658824.
XX 26-SEP-2000; 2000US-0671325.
XX 06-OCT-2000; 2000US-0677419.
XX 30-OCT-2000; 2000US-0702705.
XX 13-DEC-2000; 2000US-0736457.
XX 03-MAY-2001; 2001US-0849626.
XX (CORI-) CORIXA CORP.
XX
XX Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter WM;
XX Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS;
XX McNabb A, Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;
XX
XX WPI; 2002-164634/21.
XX P-PSDB; AAU85524.
XX
XX Novel polynucleotide encoding a lung tumour polypeptide useful for
PT stimulating and/or expanding T cells specific for a tumour protein
XX
XX Example 1; SEQ ID No 790; 223pp; English.
XX
XX The invention describes an isolated polynucleotide and polypeptide
CC useful for stimulating and/or expanding T cells specific for a tumour
CC protein for determining the presence of a cancer in a patient. A
CC composition containing the polynucleotide and/or polypeptide is useful
CC for treating a lung cancer in a patient. The polypeptide is useful for
CC removing tumour cells from a biological sample. The polynucleotide is
CC also useful as probe or primer to detect the level of mRNA encoding a
CC tumour protein. This sequence encodes a lung tumour associated protein
CC or protein fragment, described in the method of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 457 BP; 146 A; 105 C; 121 G; 85 T; 0 other;
SQ

Query Match 30.5%; Score 204; DB 24; Length 457;
Best Local Similarity 77.4%; Pred. No. 1.4e-51;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
XX 319 GTCCAGGAAACCATGTGATGAAGAGCCCTTAAGAGAGAAACACCCATTAAGTCGGAA 378
DB 120 GTCCAGGAGCCCGAGTAATGGAGAGCCGCCAAAAGAGAGACGAGCTGAAGTCGGGA 179
XX 379 TCTACACCTGATCAGAGAGAGAGATGATCAGGCTGCAGTGCAGTTCAGTTCAGTTCGA 438
DB 180 TCTACACCTGGCAGCAGACAGAGAGATCAGATCAGTGCAGTTCAGTTCAGTTCGA 239
XX 439 CTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGATGAGTGAAGGT-G 497
DB 240 CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGG 299
XX 498 GTACTGATCTCAAGGGGAGATTTCTACCAAAAGCAGAGACACTTTAAATGCCAGAGCAG 557
DB 300 GTTCCGGCTCAAGGTGAAGATAATACCTTAAGAGAGAACTGTGTAATGCCAGAGCAG 359
XX 558 GTGAAGGGAATCAGAGGTTTAAAGGAAGATAGCTGAAACCAACACAACTGTTTTATA 617
DB 360 GTGAAGAGCAACCAAGTTTAAATGAAGCAAGCTGAAACCAACGCAAGCTGGTTTTATA 419
XX 618 TTAGATATTTTACTTTAAGAGTCTTAATAATTTT 653

Db 420 TTAGATATTGACCTAAACATCTCAATAAGTTT 455

RESULT 21
ACAL1091
ID ACAL1091 standard; cDNA; 457 BP.
XX AC
XX ACAL1091;
XX
DT 05-JUN-2003 (first entry)
XX Human lung cancer-associated cDNA L52S, extended sequence.
DE Human; lung cancer; ss; lung tumour; cytostatic; vaccine;
KW T cell expansion; CD4; CD8.
XX
XX Homo sapiens.
OS
XX US2002197669-A1.
PN
XX 26-DEC-2002.
PD
XX 03-MAY-2001; 2001US-0849626.
XX
XX 13-DEC-2000; 2000US-0736457.
XX
XX (BANG/) BANGUR C S.
PA (FANG/) FANGER G R.
PA (WANG/) WANG A.
PA (WANG/) WANG T.
PA (SWIT/) SWITZER A P.
PA (MCNE/) MCNEILL P D.
PA (CLAP/) CLAPPER J D.
XX
XX Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;
PI Clapper JD;
PI
XX WPI: 2003-352750/33.
DR P-PDB; ABU69496.
XX
XX Novel lung cancer polynucleotide encoding lung cancer protein, useful
PT for detecting the presence of lung cancer in a patient, and in
PT pharmaceutical compositions, e.g. vaccines, for treating lung cancer
XX
XX Example 1; Page -: 72pp; English.
XX
XX The invention relates to a polynucleotide encoding a lung tumour protein,
CC comprising a sequence selected from any of the 14 sequences
CC mentioned in the specification, or a sequence (S2) mentioned in
CC specification, complement of S1, sequences consisting of at least 20
CC contiguous residues of S1, sequences that hybridise to S1, sequences
CC having 75% preferably 90% identity to S1, or degenerate variants of
CC S1. Also included are an isolated polypeptide (comprising a sequence (S3)
CC selected from any one of the 4 amino acid sequences mentioned in the
CC specification, a sequence encoded by the polynucleotide, or sequences
CC having at least 70%, preferably 90%, identity to a sequence encoded by
CC the polynucleotide), an expression vector comprising the polynucleotide
CC operably linked to an expression control sequence, a host cell
CC transformed or transfected with the vector, an isolated antibody (or its
CC antigen-binding fragment) that specifically binds to the polypeptide,
CC detecting the presence of a cancer in a patient, a fusion protein
CC comprising the polypeptide, an oligonucleotide that hybridises to
CC S1 under moderately stringent conditions, stimulating and/or expanding T
CC cells specific for a tumour protein (comprising contacting T cells with
CC the polynucleotide, protein or antigen-presenting cells, under conditions
CC and for a time sufficient to permit the stimulation and/or expansion of T
CC cells) and inhibiting the development of a cancer in a patient (by
CC incubating CD4⁺ and/or CD8⁺ T cells isolated from a patient with the
CC polynucleotide, protein or antigen presenting cells that express the
CC polynucleotide, such that T cells proliferate, administering to the
CC patient an effective amount of the proliferated T cells, and thus
CC inhibiting the development of a cancer in the patient. The

CC polynucleotide, protein and cells are useful in a composition for
CC stimulating an immune response in a patient, and for treating a cancer in
CC a patient (particularly lung cancer). The oligonucleotide is useful for
CC determining the presence of a cancer in a patient. The protein and
CC oligonucleotides are useful in pharmaceutical compositions, e.g. for
CC vaccines. The polynucleotide is also useful as a probe or primer for
CC nucleic acid hybridisation, and in the design and preparation of
CC ribozyme molecules for inhibiting expression of tumour polypeptides and
CC proteins in tumour cells. An amplified portion of the polynucleotide is
CC useful for isolating a full-length gene from a suitable library.
CC The present sequence is a cDNA (full length, extended or partial).
CC isolated from a library derived from lung tumour/cancer cells.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from the USPTO
CC at seqdata.uspto.gov/sequence.html?DocId=20020197669.
XX
XX Sequence 457 BP; 146 A; 105 C; 121 G; 85 T; 0 other;
SQ

Query Match 30.5%; Score 204; DB 25; Length 457;
Best Local Similarity 77.4%; Pred. No. 1.4e-51;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGGAAACCCACTGATGAAGAGCCCTAAAGAGAGAAACCCACCTAAAGTCGGAA 378
DB 120 GTCCAGGAGCCAGTAAATGGAGAGCCCCCAAAAGAGAGACCCAGCAGTGAAGTCGGGA 179
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTGCGAGCTGAGATTCAGTCCCTGA 438
DB 180 TCCTACACCTGGGAGAGAGAGATGATCAGGTGCGAGCTGAGATTCAGTCCCTGA 239
QY 439 CTTGGAAGCGGATCTCCAGGAGCTATGTCAGACAAAGACTGGGGATGGAAGGT-G 497
DB 240 CATGGAAGGTGATCTCGAGAGCTGCATCAGTCAACACCGGGGATAATCTGGATTGG 299
QY 498 GTACTGATGTCAAGGGAGAGATTTTACCAAGAGAGAGAGCATTAAATGCCAGAGCAG 557
DB 300 GTTCCGGCGCTCAGGTGAAGATAATACCTTAAAGAGAGACACTGTAAATGCCAGAGCAG 359
QY 558 GTCAAGCGGAAATCAGAGGTTTAAAGAGAGATGAAGCTGAACACAACTGTTTTATA 617
DB 360 GTGAAGAGACACCAAGTTTAAATGAAGAGAGAGCTGAACACCGAGCTGTTTATA 419
QY 618 TTAGATATTTTACTTTAAAGAGTCTTAAATTTT 653
DB 420 TTAGATATTTGACTTAAACTATCTCAATAAGTTTT 455

RESULT 22
ACA02277
ID ACA02277 standard; cDNA; 457 BP.
XX
XX ACA02277;
XX
XX 22-MAY-2003 (first entry)
DT
XX Lung cancer therapy and diagnosis associated cDNA #767.
DE Lung cancer; cytostatic; vaccine; gene therapy; cancer;
KW gene; ss.
XX
XX Homo sapiens.
OS
XX US2002172952-A1.
PN
XX 21-NOV-2002.
PD
XX 10-JUL-2001; 2001US-0302941.
PF
XX 30-JUN-1999; 99US-0346492.
PR 15-OCT-1999; 99US-0419356.
PR 17-DEC-1999; 99US-0468667.
PR 30-DEC-1999; 99US-0476300.

```
PR 06-MAR-2000; 2000US-0519642.
PR 22-MAR-2000; 2000US-0533077.
PR 10-APR-2000; 2000US-0546259.
PR 27-APR-2000; 2000US-0560406.
PR 05-JUN-2000; 2000US-0589184.
PR 11-JUL-2000; 2000US-0614124.
PR 29-AUG-2000; 2000US-0651563.
PR 08-SEP-2000; 2000US-0658824.
PR 26-SEP-2000; 2000US-0671325.
PR 06-OCT-2000; 2000US-0677419.
PR 30-OCT-2000; 2000US-0702705.
PR 13-DEC-2000; 2000US-0736457.
PR 03-MAY-2001; 2001US-0849626.
XX
XX (CORI-) CORIXA CORP.
XX
XX Henderson RA, Wang T, Watanabe Y, Johnson JC, Retter MW, Durham M;
XX Carter D, Fanger GR, Vedvick TS, Bangur CS, McNabb A;
XX
XX WPI; 2003-328427/31.
XX
XX New polynucleotide, useful for preparing a composition for treating or
XX inhibiting development of cancer, e.g. lung cancer -
XX
XX Example 1; SEQ ID NO 790; 82pp; English.
XX
XX The invention describes an isolated polynucleotide comprising one of 32
XX sequences, complement or degenerate variants of them. The polynucleotide
XX is useful for preparing a composition e.g. a vaccine or for gene therapy,
XX for treating or inhibiting development of cancer, e.g. lung cancer.
XX This sequence represents a polynucleotide associated with the
XX compositions and methods for the therapy and diagnosis of lung cancer.
XX
XX SQ Sequence 457 BP; 146 A; 105 C; 121 G; 85 T; 0 other;
XX
XX Query Match 30.5%; Score 204; DB 25; Length 457;
XX Best Local Similarity 77.4%; Pred. No. 1.4e-51;
XX Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
XX
QY 319 GTCCAGGAACCCACTGATGAGAGCCTTAAAGAGAGAAACACCCACTAAAGTCGGAA 378
DB 120 GTCCAGGAGCCAGTANTGAGAGCCCAAAAGAGAGAACACGACGTGAAGTCGGGA 179
QY 379 TCCTACACTGATCAGAGAGAGAGATGATCAGGTGCAGCTCAGATTCAGTGCCTGA 438
DB 180 TCCTACACTGGCAGCAGCAGAGAGATGATCAGGTGCAGCTCAGATTCAGTGCCTGA 239
QY 439 COTGAAGCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGATGATGTGAAGGT-G 497
DB 240 CATGAAGGTGATCTGCAAGAGCTGCATCATCAACACCGGGATAAATCTGATTGG 299
QY 498 GTACTGATGTCAGGGGAGAGATTTACCAAGCAGAGCAGCTTTAAATGCCAGAGCAG 557
DB 300 GTTCGGGCTCAAGTGAAGATAATACCTAAAGAGGACACTGTAAATGCGAGAGCAG 359
QY 558 GTGAAGGAATACAGGTTTAAAGGAGATAAGCTGAACACACAAACCTGTTTTTATA 617
DB 360 GTGAAGACACACCAAGTTTAAATGAGAGCAAGCTGAACACACGAGCTGTTTTATA 419
QY 618 TTAGATATTACTTTTAAAGAGCTTTAATAATTTT 653
DB 420 TTAGATATTGACTTAACTACTTCTCAATAAAGTTTT 455
XX
RESULT 23
XX ABA92217
XX ID ABA92217 standard; cDNA; 463 BP.
XX AC ABA92217;
XX
XX 06-JUN-2002 (first entry)
XX
XX Melanoma metastasis X-linked gene 1 (MMX-1) cDNA.
XX
```

MMX-1; melanoma metastasis X-linked gene 1; human; tumour; mammary carcinoma; breast cancer; marker; cancer/testis antigen; CTA; chromosome Xp11.21-Xp11.22; diagnosis; gene therapy; gene; ss.

Homo sapiens.

Key Location/Qualifiers
CDS 126..371
/tag= a
/product= "MMX-1_polypeptide"
polya_signal 434..439
/tag= b
primer_bind complement (48..68)
/tag= C
primer_bind 385..404
/tag= d

EP1179589-A1.
13-FEB-2002.
08-AUG-2000; 2000EP-0116253.
08-AUG-2000; 2000EP-0116253.
(HOFF) HOFFMANN LA ROCHE & CO AG F.
Not given;
WPI; 2002-229925/29.
P-PSDB; AAM51134.

New human cancer/testis antigen polypeptide, termed melanoma metastasis X-linked gene 1 polypeptide, useful as specific marker of metastatic tumor cells -

Claim 1; Fig 1; 30pp; English.

The present sequence is that of novel human melanoma metastasis X-linked gene 1 (MMX-1) cDNA. The cDNA was identified following suppression subtractive hybridisation analysis comparing common naevi and melanoma metastases. MMX-1 expression is upregulated in metastatic tumour cells and is induced during tumour expression and/or metastasis, especially in malignant melanoma and mammary carcinoma cells. MMX-1 is a specific marker of metastatic cancer cells as it can be presented in an MHC Class I complex to cytotoxic T-cells but is not presented naturally because the only non-tumour cells expressing MMX-1 are testis cells which do not present antigens in an MHC Class I context. The function of the MMX-1 gene is to promote loss of contact inhibition and anchorage dependence in tumour cells and to promote other essential steps of the metastatic cascade. Expression of the MMX-1 gene correlates with a more aggressive behaviour of the tumour cells and also with the potential for formation of metastasis. The expression profile of MMX-1 in normal tissues, tumour cell lines and tumour samples places the gene in the group of cancer/testis antigens (CTA). The MMX-1 gene has been identified, spanning about 5 kb and consisting of 4 exons and 3 introns, 2 of the introns being in the coding region. The gene is localised on the X chromosome (Xp11.21-Xp11.22) between markers DXS1204 and DXS1199, next to the CTA families of SSX, GAGE and MAGE-D. The invention provides methods for detecting the metastatic potential of melanoma and mammary carcinoma cells, and a process for determining whether a test sample of human cells has tumour progression potential, indicated by a greater amount of MMX-1 nucleic acid than cancer cell samples free of metastasis. Also provided are methods for identifying antagonists of MMX-1 or inhibitors of MMX-1 expression (e.g. antisense nucleotides) that can be used to inhibit tumour progression or metastasis and cause apoptosis of tumour cells in vivo. The MMX-1 gene and protein (see AAM51134) can also be used to identify and design drugs which interfere with the proliferation and dissemination of tumour cells.

CC comprising a targeting moiety and a detectable label, where the targeting
CC moiety specifically binds to p9p and detecting the label bound to p9p.
CC This sequence represents the XAGE-1 coding sequence that encodes XAGE-1
CC p9, an intracellular protein and the membrane-associated XAGE-1 pl6.
XX
SQ Sequence 637 BP; 196 A; 154 C; 174 G; 113 T; 0 other;

Query Match	30.5%;	Score 204;	DB 24;	Length 637;
Best Local Similarity	77.4%;	Pred. No. 1.7e-51;		
Matches 260;	Conservative	0;	Mismatches 75;	Indels 1; Gaps 1
QY	319	GTCCAGGACCCACCTGATGAAGAGCCCTAAAGAAGAAACCAACCACCTAAAGAATCGGAA	378	
DB	275	GTCCAGGAGCCAGTAATGGAGAGCCCCAAAAGAAGAACCCAGCAGCTCAAAAGTCGGA	334	
QY	379	TCCTACACCTGATCAAGAAGAGAGATGATCAGCGTGCAGCTGAGATTCAAAGTCGCTGA	438	
DB	335	TCCTACACCTGGCGCAGCAGACAGAGAAGATCAGSATACAGCTGAGATCCAGTCGCGCA	394	
QY	439	CTTGGAAAGCGGATCTCCAGGAGCTATGTGCAGACAAAGACTGGGGATGGATGTGAAGGT-G	497	
DB	395	CATGGAAAGGTGATCTGCAAGAGCTGCATCAGTCAGCAACACCGGGGATAAATCTGGATTTGG	454	
QY	498	GTACTGATGTCAAGGGGAGATTTCTACCAAAGCAGAGAGCACTTTTAAATGCCAGAAGCAG	557	
DB	455	GTTCGCGGTCAAGGTGGAAGATAATACCTTAAGAGGAAACACTGTAAATGCCAGAAGCAG	514	
QY	558	GTCAAGGGAATCACAAGGTTTAAAGGAAGATAAGCTGAAACCAACACAACCTGTTTTTATA	617	
DB	515	GTCAAGGCAACCAAGAATTTAAATGAAGACAGCTGGAACCAAGCAAGCTGGTTTTATA	574	
QY	618	TTAGATATTTTTACTTTTAAAGATCTTAATAAATTTT	653	
DB	575	TTAGATATTGACTTAAACCTACTCTCAATAAAGTTTTT	610	

RESULT 27	
AAH64751	
ID	AAH64751 standard; cDNA; 762 BP.
XX	
XX	
AC	AAH64751;
AC	
XX	
XX	
DT	11-SEP-2001 (first entry)
XX	
XX	
DE	Human secreted protein cDNA, SEQ ID NO: 27.
XX	
XX	
KW	Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW	GENSET; ss.

AA	
OS	Homo sapiens.
XX	
PN	WO200142451-A2.
XX	
PD	14-JUN-2001.
XX	
PF	07-DEC-2000; 2000WO-IB01938.
XX	
PR	08-DEC-1999; 99US-0169629.
PR	06-MAR-2000; 2000US-0187470.
XX	
PA	(GEST) GENSET.
XX	
PI	Dumas Milne Edwards J, Bougueleret L, Jobert S;
XX	
DR	WPI: 2001-367870/38.
DR	P-PSDB; AAG89148.
XX	
PT	Full length GENSET human nucleic acids encoding potentially secreted
PT	proteins, useful in gene therapy and vaccination against a variety of
PT	diseases, and for diagnosis of those diseases -
XX	
PS	Claim 7: Page 586; 9210p: English.
XX	

The invention relates to full length GENSRT human nucleic acids encoding potentially secreted proteins. The nucleic acids and the polypeptides they encode may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate GENSRT gene expression. For example, they be used to treat disorders associated with decreased GENSRT gene expression by rectifying mutations or deletions in a patient's genome that affect the activity of GENSRT or by supplementing the patients own production of GENSRT polypeptides. Conversely, antisense nucleic acid molecules may be administered to down regulate GENSRT expression by binding with the cells' own genes and preventing their expression. The sense and antisense nucleic acids may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acid sequences in samples, and hence to determine which patients may be in need of restorative therapy. The GENSRT polypeptides may be used as antigens in the production of antibodies and in assays to identify modulators (agonists and antagonists) of GENSRT polypeptide expression and activity. The present sentence is a GENSRT nucleic acid of the invention.

Sequence 762 BP; 220 A; 192 C; 186 G; 164 T; 0 other;

Query Match 30.5%; Score 204; DB 22; Length 762;
Best Local Similarity 77.4%; Pred. No. 1.8e-51;
Matches 260; Conservative 0; Mismatches 75; Indels

Qy	319	GTCCAGAGAACCCACTGATGTAAGAGCCTTAAGHAGAGAACACCACCACTATTAAGGTCCGAA	378
Db	400	GTCCAGAGGCCAGGTAAATGAGAGCCCAAAAGAGAACCCAGCAGCTGAAGATCCGGA	459
Qy	379	TCTACACCTGATCAGAAAGAGAGAGATGATCAAGGTGCAGCTGAGATCAAGTGCCCTGA	438
Db	460	TCTACACCTGGCCAGCAGACAGAGAAGATCAAGATACAGCTGAGATCCAGTGGCGCA	519
Qy	439	CCTGGNAGCCGATCTCCAGGAGCTATGTCAGACAAAGACCTGGGGATGGATGTGAAGGT-G	497
Db	520	CATGGAAGTGATCTGCAGAGAGCTGCATTCAGTCAACACC CGGGATTAATCTGGATTTGG	579
Qy	498	GTACTGATCTCAAGGGGGAAGATTCTACCAAGCAGAGCACTTTAAATGCCAAGACGAG	557
Db	580	GTTCGGCGCTCAAGGTGAAGATAATACCTTAAGAGGAACTCTGTAAATGCCAAGACGAG	639
Qy	558	GTCAAGGGAAATCAGAGCTTTTAAGGAGAGATAAGCTGAAACACACAACTGTTTTTATA	617
Db	640	GTGAAGAGCAACCAACAAGTTTAATGAAGACAGCTGAACACAGCAAGCTGTGTTTTATA	699
Qy	618	TTAGATATTTTACTTTTAAAGAGTCTTAATAAATTTT	653
Db	700	TTAGATATTTTGACTTTAAACTATCTCAATAAAGTTTTT	735

RESULT 28

AAAF68861
ID AAAF68861 standard; cDNA: 781 bp.

XX AAF68861:

XX
DT 12-APR-2001 (first entry)

Accession	Human lung tumour protein related nucleotide sequence SEQ ID NO:808.
XX	Human; lung cancer; lung tumour; lung tumour protein; gene therapy;
DE	lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccin
XX	cytotoxic; antitumor inhibition; ss.
KW	
KW	

XX Homo sapiens.

XX PN WO200100828-A2.

XX
PD
04-TAN-2001XX
DE
30-JUN-2000-

XX 30-TTN-1999: 99US-0346492

722 TTAGATATTGACTTAACTATCTCAATAAAGTTTT 757

Db

RESULT 29
ABK38772

ID ABK38772 standard; cDNA; 781 BP.
XX
AC ABK38772;
XX
21-MAY-2002 (first entry)
DT
DE cDNA encoding L552S lung tumour protein.
XX
XX Lung tumour; cancer; T cell; immune response stimulator;
KW cytestatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200204514-A2.
XX
XX 17-JAN-2002.
XX
XX 10-JUL-2001; 2001WO-US22058.
XX
XX 11-JUL-2000; 2000US-0614124.
XX 29-AUG-2000; 2000US-0651563.
XX 08-SEP-2000; 2000US-0658824.
XX 26-SEP-2000; 2000US-0671325.
XX 06-OCT-2000; 2000US-0677419.
XX 30-OCT-2000; 2000US-0702705.
XX 13-DEC-2000; 2000US-0736457.
XX 03-MAY-2001; 2001US-0849626.
XX
XX (CORI-) CORIXA CORP.
XX
XX Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;
PI Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS;
PI McNabb A, Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;
XX
XX WPI: 2002-164634/21.
XX P-PSDB; AAU85651.
XX
XX Novel polynucleotide encoding a lung tumour polypeptide useful for
PT stimulating and/or expanding T cells specific for a tumour protein -
XX
XX Example 1; SEQ ID No 808; 223pp; English.
XX
XX The invention describes an isolated polynucleotide and polypeptide
XX useful for stimulating and/or expanding T cells specific for a tumour
XX protein for determining the presence of a cancer in a patient. A
XX composition containing the polynucleotide and/or polypeptide is useful
XX for treating a lung cancer in a patient. The polypeptide is useful for
XX removing tumour cells from a biological sample. The polynucleotide is
XX also useful as probe or primer to detect the level of mRNA encoding a
XX tumour protein. This sequence encodes a lung tumour associated protein
XX or protein fragment, described in the method of the invention.
XX Note: the sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 781 BP; 230 A; 181 C; 234 G; 136 T; 0 other;
SQ
Query Match 30.5%; Score 204; DB 24; Length 781;
Best Local Similarity 77.4%; Pred. No. 1.8e-51;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGAGACCCACTGATGAAGAGGCTTAAGAAAGAGAAACACCCACTAAAGTCGGAA 378
DB 422 GTCCAGAGAGCCCAAGTAAATGAGAGGCCCAAAAAGAAAGAACCCAGAGCTGAAGTCGGGA 481
QY 379 TCTTACCTCATCAGAAAGAGAGAAATGATCAGGTGCAGCTGAGATTCAAGTGCCTGA 438

```
Db 482 TCCTACACCTGGCGCAGCAGACAGAGAAGATCAGGATACAGCTGAGATCCAGTGGCGGA 541
Qy 439 CCTGGAGCGGATCTCCAGGAGCTATGTTCAGACAAAGACTGGGATGGATGTGAAGT-G 497
Db 542 CATGGAAGGTGATCTGAGAGCTGCATCAGTCAACACCGGGATAAATCTGATTTGG 601
Qy 498 GTACTGATGTCAAGGGGAAGATTCTTACAAAGCAGAGCAGCTTTAAATGCCAGAGCAG 557
Db 602 GTTCGGCGTCAAGGTGAAGATAATACCTAAAGAGGAACACTGTAAATGCCAGAGCAG 661
Qy 558 GTGAAGGAAATCACAGGTTTAAAGGAGATAAGCTGAACAAACACAACTGTTTTTATA 617
Db 662 GTGAAGACACACCAAGCTTAAATGAAGACAGCTGAACACGCAAGCTGGTTTTATA 721
Qy 618 TTAGATATTACTTTAAAGAGCTTTAAATAATTTT 653
Db 722 TTAGATATTGACTTAAACTATCTCATATAAGTTTT 757

RESULT 30
ACAL1101
ID ACAL1101 standard; cDNA; 781 BP.
XX AC
XX ACAL1101;
XX
DT 05-JUN-2003 (first entry)
XX
DE Human lung cancer-associated full length cDNA L5528.
XX
XX Human; lung cancer; ss; lung tumour; cytostatic; vaccine;
XX T cell expansion; CD4; CD8.
XX
XX Homo sapiens.
XX
XX US2002197669-A1.
XX
XX 26-DEC-2002.
XX
XX 03-MAY-2001; 2001US-0849626.
XX
XX 13-DEC-2000; 2000US-0736457.
XX
XX (BANG/) BANGUR C S.
XX (FANG/) FANGER G R.
XX (WANG/) WANG A.
XX (WANG/) WANG T.
XX (SWIT/) SWITZER A P.
XX (MCNE/) MCNEILL P D.
XX (CLAP/) CLAPPER J D.
XX
XX Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;
XX Clapper JD;
XX
XX MPI; 2003-352750/33.
XX P-PSDB; ABU69504.
XX
XX Novel lung cancer polynucleotide encoding lung cancer protein, useful
XX for detecting the presence of lung cancer in a patient, and in
XX pharmaceutical compositions, e.g. vaccines, for treating lung cancer -
XX
XX Example 1; Page -; 72pp; English.
XX
XX The invention relates to a polynucleotide encoding a lung tumour protein,
XX comprising a sequence selected from any of the 14 sequences
XX mentioned in the specification, or a sequence (S2) mentioned in
XX the specification, complement of S1, sequences consisting of at least 20
XX contiguous residues of S1, sequences that hybridise to S1, sequences
XX having 75%, preferably 90%, identity to S1, or degenerate variants of
XX S1. Also included are an isolated polypeptide (comprising a sequence (S3)
XX selected from any one of the 4 amino acid sequences mentioned in the
XX specification, a sequence encoded by the polynucleotide, or sequences
XX having at least 70%, preferably 90%, identity to a sequence encoded by
XX the polynucleotide), an expression vector comprising the polynucleotide
```

```
CC operably linked to an expression control sequence, a host cell
CC transformed or transfected with the vector, an isolated antibody (or its
CC antigen-binding fragment) that specifically binds to the polypeptide,
CC detecting the presence of a cancer in a patient, a fusion protein
CC comprising the polypeptide, an oligonucleotide that hybridises to
CC si under moderately stringent conditions, stimulating and/or expanding T
CC cells specific for a tumour protein (comprising contacting T cells with
CC the polynucleotide, protein or antigen-presenting cells, under conditions
CC and for a time sufficient to permit the stimulation and/or expansion of T
CC cells) and inhibiting the development of a cancer in a patient (by
CC incubating CD4+ and/or CD8+ T cells isolated from a patient with the
CC polynucleotide, protein or antigen presenting cells that express the
CC polynucleotide, such that T cells proliferate, administering to the
CC patient an effective amount of the proliferated T cells, and thus
CC inhibiting the development of a cancer in the patient. The
CC polynucleotide, protein and cells are useful in a composition for
CC stimulating an immune response in a patient, and for treating a cancer in
CC a patient (particularly lung cancer). The oligonucleotide is useful for
CC determining the presence of a cancer in a patient. The protein and
CC oligonucleotides are useful in pharmaceutical compositions, e.g.
CC vaccines. The polynucleotide is also useful as a probe or primer for
CC nucleic acid hybridisation, and in the design and preparation of
CC ribozyme molecules for inhibiting expression of tumour polypeptides and
CC proteins in tumour cells. An amplified portion of the polynucleotide is
CC useful for isolating a full-length gene from a suitable library.
CC The present sequence is a cDNA (full length, extended or partial)
CC isolated from a library derived from lung tumour/cancer cells.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from the USPTO
CC at seqdata.uspto.gov/sequence.html?DocId=20020197669.
XX
XX Sequence 781 BP; 230 A; 181 C; 234 G; 136 T; 0 other;
SQ
Query Match 30.5%; Score 204; DB 25; Length 781;
Best Local Similarity 77.4%; Pred. No. 1.8e-51;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
Qy 319 GTCCGAGGAAACCACCTGATGAGAGAGCTAAAGAGAGAAACCACCTAAAGTCGGAA 378
Db 422 GTCCGAGGAGCCCAAGTAATGGAGAGCCCAAAAGAGAGAAACCACCTGAAAGTCGGAA 481
Qy 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCGAGCTGAGATTCAGATTCCTGA 438
Db 482 TCCTACACCTGGCAGCAGACAGAGAGATCAGGATACAGCTGAGATCCAGTGGCGA 541
Qy 439 CTTGGAAGCCGATCTCCAGAGCTATGTTCAGACAAAGACTGGGGATGGATGTGAAGT-G 497
Db 542 CATGGAAGGTGATCTGCAAGAGCTGATCAGTCAAAACACCGGGGATAAATCTGGATTTGG 601
Qy 498 GTACTGATGTCAAGGGGAAGATTCTACAAAGCAGAGCAGACTTTAAATGCCAGAGCAG 557
Db 602 GTTCCGCGCTCAAGGTGAAGATAATACCTAAAGAGGACACTGTAAATGCCAGAGCAG 661
Qy 558 GTGAAGGGAATCACAGGTTTAAAGGAGATAAGCTGAACAAACACAACTGTTTTTATA 617
Db 662 GTGAAGACACACCAAGTTTAAATGAAGACAGCTGAACAAACGCAAGCTGGTTTTATA 721
Qy 618 TTAGATATTACTTTAAAGAGCTTTAAATAATTTT 653
Db 722 TTAGATATTGACTTAAACTATCTCATATAAGTTTT 757

RESULT 31
ACA02287
ID ACA02287 standard; cDNA; 781 BP.
XX AC
XX ACA02287;
XX
DT 22-MAY-2003 (first entry)
XX
XX Lung cancer therapy and diagnosis associated cDNA #776.
XX
```



```
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 750 BP; 225 A; 164 C; 203 G; 158 T; 0 other;

  Query Match      27.5%; Score 193.8; DB 23; Length 750;
  Best Local Similarity 86.4%; Pred. No. 2.5e-45;
  Matches 203; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 323 CAGNACCCACTGATGAGAGCCTTAAGAAGAGAAACACCCACTAAAGTCGGAATCCT 382
Dd |||||
QY 208 CAGAGCCCATGTGCGCAGAGCCTCAACAGAAAGAACCAACCACTGAAGATCAGGATCAT 267
Dd |||||
QY 383 ACACCTGATCAGAAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTSCCTGACCTG 442
Dd |||||
QY 268 ACACCTGCTCAGAAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTSCCTTAACCTG 327
Dd |||||
QY 443 GAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGATGATGAGGTGGTACT 502
Dd |||||
QY 328 GAGGCTGATCTCCAGGAGCTGTCTCAGTCAAGACTGGGGATGAATCGCGAGATAGTCCT 387
Dd |||||
QY 503 GATGTCAGAGGGGAGATTCTACCAAAAGCAGAGACATTTAAATGCCAGAGCAG 557
Dd |||||
QY 388 GATGTCAGAGGGGAGATTCTCCAAATCAGAGCAATTTAAATGCCAGAGGAG 442
Dd |||||

RESULT 35
ABZ20463
ID ABZ20463 standard; cDNA; 530 BP.
XX
AC ABZ20463;
XX
DT 23-JAN-2003 (first entry)
XX
DE GAGE-2 full length cDNA sequence SEQ ID NO:2890.
XX
KW Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;
KW immune response; virology; immunology; microbiology; molecular biology;
KW recombinant DNA technology; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200278516-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US10421.
XX
PR 30-MAR-2001; 2001US-280255P.
XX
PR 28-AUG-2001; 2001US-315563P.
XX
PR 09-JAN-2002; 2002US-347313P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Wang S, Bangur CS, Gaiger A;
XX
DR WPI; 2003-058387/05.
XX
DR P-PSDB; ABP54447.
XX
PT New immunogenic polynucleotides or polypeptides useful for diagnosing,
PT preventing and treating cancer expressing CT or CP mRNA antigens, and
PT in virology, immunology, microbiology, molecular biology and
PT recombinant DNA techniques
XX
PS Claim 1; SEQ ID 2890; 207pp; English.
XX
CC ABQ17575 to ABQ20506 represent isolated polynucleotide (i) sequences, and
CC ABP54446 to ABP54472 represent protein (ii) sequences, from the present
CC invention. (i) and (ii) have cytostatic activity and can be used in gene
CC therapy and vaccines. (i), (ii), antibodies and compositions from the
CC present invention are useful for diagnosing, preventing and treating
```

```
CC cancer, which expresses CT or CP mRNA antigens. They are useful for
CC stimulating immune response. They can also be useful in virology,
CC immunology, microbiology, molecular biology and recombinant DNA
CC techniques.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 530 BP; 152 A; 116 C; 155 G; 107 T; 0 other;

  Query Match      27.1%; Score 181.2; DB 25; Length 530;
  Best Local Similarity 67.4%; Pred. No. 1.3e-44;
  Matches 287; Conservative 0; Mismatches 133; Indels 6; Gaps 2;

QY 217 GTGAATATGATGTGGCGAGGAAGATCAACATATATAGGCTAGGCCAAGAGAGTTTACA 276
Dd |||||
Dd 77 GTGAATATGATGTGGCGAGGAAGATCGACCTATCGGCTAGACCACAGACGCTACGTAGA 136
Dd |||||
QY 277 GCCTCCTCAGCTGATTTGGGGCTATGCTTACTGGCTCCCTTTGTCCCGAGAACCCACTGA 336
Dd |||||
Dd 137 GCCTCCTGAAATGATTTGGGCTATGGGCC---CGAGCAGTTTCAGTGAAGTGAACC 193
Dd |||||
QY 337 TGAAGAGCCTTAAGAGAGAGAAACCCACCCACTAAAGTGGAAATCCTACCTGATCAGNA 396
Dd |||||
Dd 194 AGCAACACCTGAAGAGAGGGGAAACCCAGCACTCAAGTCAGGATCCTGCAGCTGCTCAGGA 253
Dd |||||
QY 397 GAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAGCCGATCTCCA 456
Dd |||||
Dd 254 GGGAGAGGATG---AGGGAGCATCTGAGGTCAAGGCCGAGCTGAAGCTCATAGCCA 310
QY 457 GGAGCTATCTCAGACAAAGACTGGGGATGGATGTGAAGTGGTGTACTGTATGTCAAGGGHAA 516
Dd |||||
Dd 311 GGAACAGGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTGTATGGCAGGAGAT 370
QY 517 GATTCTACCAAGAGCAGACACTTTAAATGCCAGAGCAGCTGAAGGGAATCAGGT 576
Dd |||||
Dd 371 GAGCCGCCAATCTCCAGAGAGGTGAAGCCCTGAAGAGGTGAAGCAATCAGAGTG 430
QY 577 TTAAGAGAGATAAGCTGAACAAACACAACTGTTTTATATAGATATTTTACTTTAAA 636
Dd |||||
Dd 431 TTAAGAGAGACACGTTGAATGATGACAGGCTGCTCTATGTTGGAATTTTGTTCATTAA 490
QY 637 GAGTCT 642
Dd 491 AATTCT 496

RESULT 36
AAV18717
ID AAV18717 standard; cDNA; 535 BP.
XX
AC AAV18717;
XX
DT 30-JUL-1998 (first entry)
XX
DE cDNA encoding GAGE-2 tumour rejection antigen precursor.
XX
KW GAGE tumour rejection antigen precursor; TRAP; tumour;
KW diagnosis; melanoma; antigen; cytolytic T cell clone proliferation;
KW HLA-typing assay; ss.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
XX CDS 81..431
XX /*tag= a
XX /transl_except= (pos:192..194, aa:Ala)
XX /transl_except= (pos:195..197, aa:Thr)
XX
XX W09749417-A1.
XX
XX 31-DEC-1997.
XX
```

```

PF 23-JUN-1997; 97WO-US10850.
XX
XX PR 24-JUN-1996; 96US-0669161.
XX
XX PA (LUDW-) LUDWIG INST CANCER RES.
XX
XX PI Boon-Falleur T, Debacker O, Van Den Eynde B;
XX
XX DR WPI; 1998-076905/07.
XX
XX DR P-PSDB; AAW47599.
XX
XX PT Isolated nucleic acid encoding GAGE tumour rejection antigen
XX precursor - processed by HLA-Cw6 molecules into peptides, useful to
XX diagnose melanomas
XX
XX PS Example 13; Fig 4; 60pp; English.
XX
XX CC The present sequence encodes a GAGE-2 tumour rejection antigen
XX precursor (TRAP). The protein is expressed in a number of tumours. In
XX contrast the only normal tissue which expresses GAGE TRAP protein is
XX testis. Several GAGE TRAPs have been identified (see AAV18717-21). The
XX major difference between these proteins and GAGE-1 (AAV05540) is the
XX absence of a stretch of 143 bases located at position 379 to 521 of the
XX GAGE-1 TRAP sequence. The rest of the sequences show mismatches at
XX various position, with the exception of GAGE-3 whose 5' end is totally
XX different from the other GAGE cDNAs for the first 112 bases. This
XX region of GAGE-3 cDNA contains a long repeat and a hairpin structure.
XX The antigens can be used to diagnose melanomas, characterised by
XX expression of a TRAP or presentation of a tumour rejection antigen.
XX Antigens shed into blood or urine can be observed and then used to
XX confirm a diagnosis of melanoma using cytolytic T cell clone
XX proliferation methodologies. Other uses for the processed peptides,
XX include HLA-typing assays for, e.g. skin graft or organ transplants.
XX
XX SQ Sequence 535 BP; 158 A; 112 C; 156 G; 109 T; 0 other;

Query Match 27.1%; Score 181.2; DB 19; Length 535;
Best Local Similarity 67.4%; Pred. No. 1.3e-44;
Matches 287; Conservative 0; Mismatches 133; Indels 6; Gaps 2;

QY 217 GTGAAATATGAGTTCGAGGAGAGATCAACATATAGGCTAGGCCAAGAAAGTTTACA 276
DB 74 GTGAAATATGAGTTCGAGGAGAGATCGACCTATCGGCTAGCAAGACGCTACGTAGA 133

QY 277 GCCTCTGAGCTGATTGGGGCTATGCTTACTGGCTCCCTTTGTCCAGGACCCACTGA 336
DB 134 GCCTCTGAAATGATTGGGCTATGCGGCC---CGAGCAGTTTCAGTGATGAGTGGAAACC 190

QY 337 TGAAGAGCCTAAAGAGAGAGAAACCCACCTAAAGTCGGAATCCTACACCTGATCAGAA 396
DB 191 AGAGCCTCTGAAAGAGGGGAAACCCAGCAACTCAAGTCAGGATCCTGCAGCTGCTCAGGA 250

QY 397 GAGAGAGATGATCAGGTTGCGAGTTCAGATTCAGTGCTGACCTGGAGCGGATCTCCA 456
DB 251 GGGAGAGGATG---AGGGAGCATCTCAGGTCAAGGGCCGAAAGCCTGAAGCTGAGAGCCA 307

QY 457 GGAGCTATGTCAGACAAAGATGGGGATGGATGTGAAGTGGTACTGATGTCAAGGGGAA 516
DB 308 GGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGTCCTGATGGCCAGGAGAT 367

QY 517 GATTCTACCAAGCAGAGCACTTTAAATGCCAGNAGCAGGTGAAGGAATTCAGGT 576
DB 368 GGACCCGCCAAATCCAGAGAGGATGAAGACCGCTGAAGAGGTGAAGCAATCAGTG 427

QY 577 TTAAGGAAGATAGCTGAAACAAACAACTGTTTTATATTAGATATTTTACTTTAAA 636
DB 428 TTAAGGAAGACATGTTGAAATGATGACGGCTGCTCTATGTTGGAATTTGTCATTAA 487

QY 637 GAGTCT 642
DB 488 AATTCT 493

```

```

RESULT 37
AAK90519
ID AAK90519 standard; cDNA; 538 BP.
XX
XX AC AAK90519;
XX
XX DT 30-SEP-1999 (first entry)
XX
XX DE GAGE-2 tumour rejection antigen clone nucleotide sequence.
XX
XX KW Human leukocyte antigen; HLA-A29; tumour rejection antigen;
XX detection; therapy; pathological condition; cancer; CTL;
XX cytolytic T lymphocyte; GAGE; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO9937665-A1.
XX
XX PD 29-JUL-1999.
XX
XX PF 12-JAN-1999; 99WO-US00775.
XX
XX PR 23-JAN-1998; 98US-0012818.
XX
XX PA (LUDW-) LUDWIG INST CANCER RES.
XX
XX PI Boon-Falleur T, Debacker O, Van Den Eynde B, Van Der Bruggen P;
XX WPI; 1999-469111/39.
XX
XX PT New isolated peptides which bind to HLA-A29 molecules, which are
XX tumour rejection antigens used for detection and therapy of
XX pathological conditions, e.g. cancer
XX
XX PS Example 13; Fig 4; 62pp; English.
XX
XX CC The present invention describes peptides which bind to human leukocyte
XX antigen (HLA) A29 (HLA-A29) molecules. The peptides are processed into
XX tumour rejection antigens. They can be used for detecting cytolytic T
XX lymphocytes (CTLs) in pathological conditions such as cancer and in
XX HLA-typing assays. Complexes of HLA-A29 molecules and the peptides can
XX be used for stimulating CTLs in vivo. The present sequence represents
XX a GAGE tumour rejection antigen clone, from an example from the present
XX invention.
XX
XX SQ Sequence 538 BP; 160 A; 116 C; 155 G; 107 T; 0 other;

Query Match 27.1%; Score 181.2; DB 20; Length 538;
Best Local Similarity 67.4%; Pred. No. 1.3e-44;
Matches 287; Conservative 0; Mismatches 133; Indels 6; Gaps 2;

QY 217 GTGAAATATGAGTTCGAGGAGAGATCAACATATAGGCTAGGCCAAGAAAGTTTACA 276
DB 77 GTGAAATATGAGTTCGAGGAGAGATCGACCTATCGGCTAGCAAGACGCTACGTAGA 136

QY 277 GCCTCTGAGCTGATTGGGGCTATGCTTACTGGCTCCCTTTGTCCAGAAACCCACTGA 336
DB 137 GCCTCTGAAATGATTGGGCTATGCGGCC---CGAGCAGTTTCAGTGATGGAAGTGAAC 193

QY 337 TGAAGAGCCTAAAGAGAGAAACCCACCTAAAGTCGGAATCCTACACCTGATCAGAA 396
DB 194 AGCAACCTTGAAAGAGGGGAAACCACTCAAGTTCAGGATCTTCGAGCTGCTCAGGA 253

QY 397 GAGAGAGATGATCAGGTTGCGAGTTCAGATTCAGTTCCTGAGCTCGAAGCCGATCTCA 456
DB 254 GGGAGAGGATG---AGGGAGCATCTGCGGTCAAGGGCCGAAAGCCTGAAGCTCATAGCCA 310

QY 457 GGAGCTATGTCAGACAAAGATGGGGATGGATGTGAAGTGGTACTGATGTCAAGGGGAA 516
DB 311 GGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGTGGTTCCTGATGGGAGGAGAT 370

QY 517 GATTCTACCAAGCAGAGCACTTTAAATGCCAGAGCAGGTGAAGGAAATCAGAGT 576

```


Db 371 GGACCGCCAAATCCAGAGGAGTGAACCCCTGAAGAGGTGAAGAAGCAATACACAGTG 430
Qy 577 TTAAGGAGATAGCTGAACACACAAACTGTTTATATATAGATATTACTTTAA 636
Db 431 TTAAGAAGACACGTTGAATGATGACAGGCTGCTCTATGTTGAAATTTGTTCAAT 490
Qy 637 GAGTCT 642
Db 491 AATTCT 496
RESULT 38
AAF68152
ID AAF68152 standard; cDNA; 479 BP.
XX AC AAF68152;
XX DT 12-APR-2001 (first entry)
XX DE Human lung tumour protein related nucleotide sequence SEQ ID NO:70.
XX KW Human; lung cancer; lung tumour; lung tumour protein; gene therapy;
KW lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;
KW cytostatic; antisense inhibition; ss.
XX OS Homo sapiens.
XX FN WO200100828-A2.
XX PD 04-JAN-2001.
XX PF 30-JUN-2000; 2000WO-US18061.
XX PR 30-JUN-1999; 99US-0346492.
PR 15-OCT-1999; 99US-0419356.
PR 17-DEC-1999; 99US-0456867.
PR 30-DEC-1999; 99US-0476300.
PR 06-MAR-2000; 2000US-0519642.
PR 22-MAR-2000; 2000US-0532077.
PR 10-APR-2000; 2000US-0548259.
PR 27-APR-2000; 2000US-0560406.
PR 05-JUN-2000; 2000US-0589184.
XX PA (CORI-) CORIXA CORP.
XX PI Wang T, Bangur CS, Lodes MJ, Fanger GR, Vedvick TS, Carter D;
PI Retter MW, Mannion J;
XX WPI; 2001-071488/08.
XX PT Lung tumor-associated proteins and the nucleic acids that encode them,
XX useful for preventing, diagnosing and treating lung cancer -
XX PS Claim 4; Page 174; 436pp; English.
XX CC The present invention describes immunogenic portions of lung tumour-
CC associated proteins (I) and the nucleic acids (NAs) that encode them.
CC (I) have cytostatic activity and can be used in gene therapy, antisense
CC inhibition and in vaccines. The NAs and the lung tumour-associated
CC proteins they encode may be used in the prevention, treatment and
CC diagnosis of diseases associated with their inappropriate expression,
CC especially lung cancers. For example, the NAs may be administered to
CC treat diseases by rectifying mutations or deletions in a patient's genome
CC that affect the activity of the protein by expressing inactive proteins
CC or to supplement the patients own production of (I). Additionally, the
CC NAs may be used to produce the lung-tumour associated protein, according
CC to standard recombinant DNA methodology. Conversely, antisense NA
CC molecules may be administered to down regulate protein expression by
CC binding with the cells own genes and preventing their expression. The NA
CC and complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar NA sequences in
CC samples, and hence which patients may be in need of treatment for lung
CC cancer. The (I) may be used as antigens in the production of antibodies

CC and in assays to identify modulators (agonists and antagonists) of the
CC expression and activity of the protein. AAF68083 to AAF68878 and
CC AAB76848 to AAB76878 represent human lung tumour protein related
CC nucleotide and protein sequences which are used in the exemplification
CC of the present invention.
XX SQ Sequence 479 BP; 163 A; 107 C; 125 G; 84 T; 0 other;
Query Match 26.9%; Score 180; DB 22; Length 479;
Best Local Similarity 74.0%; Pred. No. 2.8e-44;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
Qy 319 GTCCAGGAACCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCTATAAAGTCGGAA 378
Db 103 GTCCAGGAGCCAGTATGGAGAGCCCCCAAAAGAAAGAACAGCAGCTGAAGTCGGGA 162
Qy 379 TCTACACCTGATCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTC----- 428
Db 163 TCTACACCTGGGCGACAGACAGAAAGATCAGGATACAGCTGAGATCCCGAGTCTG 222
Qy 429 -----AAGTCTGACCTGGAGGCGGATCTCCAGAGCTATGTCAGACAAAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAAGTGATCTCCAAAGAGCTGCATCAGTCAACACCGGGG 282
Qy 483 ATGGATGTGAAGGT-GGTACTGATGTCAAGGGGAAGATTCTACCAAAAGCAGAGCACTTT 541
Db 283 ATAAATCTGGATTGGGTTCCGCGCTCAAGGTGAAGATAATACCTAAAGAGGAACACTGT 342
Qy 542 AAAATGCCAGAGCAGGTGAAGGGAATCACAGGTTTAAAGGAAGATAAGCTGAACACAC 601
Db 343 AAAATGCCAGAGCAGGTGAAGGGAATCACAGGTTTAAAGGAAGATAAGCTGAACACAC 402
Qy 602 ACAAACTGTTTTATATTAGATATTCTTTAAAGAGTCTTTAATAAATTTTG 655
Db 403 GCAAGCTGTTTTATATTAGGATATTGACTTTAAACTATCTCAATAAAGTTTG 456
RESULT 39
ABK38063
ID ABK38063 standard; cDNA; 479 BP.
XX AC ABK38063;
XX DT 21-MAY-2002 (first entry)
XX DE cDNA encoding clone #19106 (L547S) of lung tumour protein.
XX KW Lung tumour; cancer; T cell; immune response stimulator;
XX KW cytostatic; gene; ss.
XX OS Homo sapiens.
XX PF WO200204514-A2.
XX PD 17-JAN-2002.
XX PR 10-JUL-2001; 2001WO-US22058.
XX PR 11-JUL-2000; 2000US-0614124.
PR 29-AUG-2000; 2000US-0651563.
PR 08-SEP-2000; 2000US-0658824.
PR 26-SEP-2000; 2000US-0671325.
PR 06-OCT-2000; 2000US-0677419.
PR 30-OCT-2000; 2000US-0703705.
PR 13-DEC-2000; 2000US-0736457.
PR 03-MAY-2001; 2001US-0849626.
XX PA (CORI-) CORIXA CORP.
XX PI Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;
PI Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS;
PI McNabb A, Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;
XX

DR WPI; 2002-164634/21.
XX Novel polynucleotide encoding a lung tumour polypeptide useful for
PT stimulating and/or expanding T cells specific for a tumour protein
XX Example 1; SEQ ID No 70; 223pp; English.
XX The invention describes an isolated polynucleotide and polypeptide
CC useful for stimulating and/or expanding T cells specific for a tumour
CC protein for determining the presence of a cancer in a patient. A
CC composition containing the polynucleotide and/or polypeptide is useful
CC for treating a lung cancer in a patient. The polypeptide is useful for
CC removing tumour cells from a biological sample. The polynucleotide is
CC also useful as probe or primer to detect the level of mRNA encoding a
CC tumour protein. This sequence encodes a lung tumour associated protein
CC or protein fragment, described in the method of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 479 BP; 163 A; 107 C; 125 G; 84 T; 0 other;
XX
XX Query Match 26.9%; Score 180; DB 24; Length 479;
XX Best Local Similarity 74.0%; Pred. No. 2.8e-44;
XX Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
XX
XX 319 GTCCAGAACCCCTGATGAAGAGCCTAAAGAGAGAAACACCCACTAAAGTCCGAA 378
XX 103 GTCCAGAGCCCACTAATGGAGCCCCAAAGAGAACACGACGCTGAATCGGGA 162
XX 379 TCTACACCTGATCAGAGAGAGAGATGATCAGGGTCCAGTGCAGATTC----- 428
XX 163 TCTACACCTGGCAGCAGACAGAGAGATCAGGATACAGTGCAGATTCCTCCAGTGG 222
XX 429 -----AAGTGCCTGACCTGGAGCGGATCTCCAGAGCTATGTCAGACAAAGACTGGG 482
XX 223 GAAGGGAATGCGGACATGGAGGTGATCTCAAGAGCTGATCAGTCAACACCGGG 282
XX 483 ATGAGTGAAGGT-GGTACTGATGTCAGGGGAAGATTCACCAAAAGCAGACACTTT 541
XX 283 ATAAATCTGGATTGGGTTCCGGGCTCAAGGTGAAGATAAATACCTAAAGAGGAACACTGT 342
XX 542 AAATGCCAGAGACGGTGAAGGAAATCACAGGTTTAAAGAAAGATGAAGTGAACAC 601
XX 343 AAATGCCAGAGACGGTGAAGAGCAACCAAGTTTAAATGAAGACAGCTGAACAC 402
XX 602 ACAAACTGTTTTATATAGATATTTTACTTTAAAGAGTCTTTAATAAATTTTG 655
XX 403 GCAAGCTGGTTTTATATAGGATATTTGACTTAAACTTCTCAATAAAGTTTGG 456
XX
XX RESULT 40
XX ID ACA10392
XX AC ACA10392;
XX XX
XX DT 05-JUN-2003 (first entry)
XX DE Human lung cancer-associated cDNA, L547D.
XX KW Human; lung cancer; ss; lung tumour; cytostatic; vaccine;
XX T cell expansion; CD4; CD8.
XX OS Homo sapiens.
XX XX
XX XX US2002197669-A1.
XX PD 26-DEC-2002.
XX XX
XX PF 03-MAY-2001; 2001US-0849626.
XX XX

PR 13-DEC-2000; 2000US-0736457.
XX (BANG/) BANGUR C S.
PA (FANG/) FANGER G R.
PA (WANG/) WANG A.
PA (WANG/) WANG T.
PA (SWIT/) SWITZER A P.
PA (MCNE/) MCNEILL P D.
PA (CLAP/) CLAPPER J D.
XX Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;
PI Clapper JD;
PI WPI; 2003-352750/33.
XX Novel lung cancer polynucleotide encoding lung cancer protein, useful
PT for detecting the presence of lung cancer in a patient, and in
PT pharmaceutical compositions, e.g. vaccines, for treating lung cancer
XX
XX Example 1; Page -; 72pp; English.
XX The invention relates to a polynucleotide encoding a lung tumour protein,
XX comprising a sequence selected from any of the 14 sequences
XX mentioned in the specification, or a sequence (S2) mentioned in
XX specification, complement of S1, sequences consisting of at least 20
XX contiguous residues of S1, sequences that hybridise to S1, sequences
XX having 75%, preferably 90%, identity to S1, or degenerate variants of
XX S1. Also included are an isolated polypeptide (comprising a sequence (S3)
XX selected from any one of the 4 amino acid sequences mentioned in the
XX specification, a sequence encoded by the polynucleotide, or sequences
XX having at least 70%, preferably 90%, identity to a sequence encoded by
XX the polynucleotide), an expression vector comprising the polynucleotide
XX operably linked to an expression control sequence, a host cell
XX transformed or transfected with the vector, an isolated antibody (or its
XX antigen-binding fragment) that specifically binds to the polypeptide,
XX detecting the presence of a cancer in a patient, a fusion protein
XX comprising the polypeptide, an oligonucleotide that hybridises to
XX S1 under moderately stringent conditions, stimulating and/or expanding T
XX cells specific for a tumour protein (comprising contacting T cells with
XX the polynucleotide, protein or antigen-presenting cells, under conditions
XX and for a time sufficient to permit the stimulation and/or expansion of T
XX cells) and inhibiting the development of a cancer in a patient (by
XX incubating CD4+ and/or CD8+ T cells isolated from a patient with the
XX polynucleotide, protein or antigen presenting cells that express the
XX polynucleotide, such that T cells proliferate, administering to the
XX patient an effective amount of the proliferated T cells, and thus
XX inhibiting the development of a cancer in the patient. The
XX polynucleotide, protein and cells are useful in a composition for
XX stimulating an immune response in a patient, and for treating a cancer in
XX a patient (particularly lung cancer). The oligonucleotide is useful for
XX determining the presence of a cancer in a patient. The protein and
XX oligonucleotides are useful in pharmaceutical compositions, e.g.
XX vaccines. The polynucleotide is also useful as a probe or primer for
XX nucleic acid hybridisation, and in the design and preparation of
XX ribozyme molecules for inhibiting expression of tumour polypeptides and
XX proteins in tumour cells. An amplified portion of the polynucleotide is
XX useful for isolating a full-length gene from a suitable library.
XX The present sequence is a cDNA (full length, extended or partial)
XX isolated from a library derived from lung tumour/cancer cells.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from the USPTO
XX at seqdata.uspto.gov/sequence.html?DocId=20020197669.
XX
XX Sequence 479 BP; 163 A; 107 C; 125 G; 84 T; 0 other;
XX
XX Query Match 26.9%; Score 180; DB 25; Length 479;
XX Best Local Similarity 74.0%; Pred. No. 2.8e-44;
XX Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
XX
XX 319 GTCCAGAACCCCTGATGAAGAGCCTAAAGAGAGAAACACCCACTAAAGTCCGAA 378
XX 103 GTCCAGAGCCCACTAATGGAGCCCCAAAGAGAACACGACGCTGAATCGGGA 162
XX

XX The invention relates to 1046 novel nucleic acids which are used as
 CC markers for determining the sensitivity of a cancer cell to the
 CC anticancer agent TAXOL. Cancer cells can be treated with TAXOL when
 CC they are shown to express one of the 242 sensitivity markers or the
 CC cells are shown not to express one of the 804 resistance markers.
 CC The methods can be used to determine the effectiveness of TAXOL
 CC in the treatment of cancer cell growth in an individual. The markers
 CC can be used as targets in developing anti-cancer agents such as
 CC chemotherapeutic compounds. The markers can also be used as targets in
 CC developing treatments for cancer, particularly those cancers which
 CC display resistance to agents and exhibit expression of the markers. The
 CC anticancer agents developed by the novel method can be used to treat
 CC cancer. Probes based on the markers can be used to detect transcripts or
 CC genomic sequences corresponding to the markers, in the identification of
 CC cells or tissues which mis-express the protein. Cancers which may
 CC be targeted include carcinoma (e.g. squamous cell carcinoma),
 CC sarcoma (e.g. fibrosarcoma) leukemia (e.g. lymphocytic leukemia),
 CC lymphoma (e.g. plasmacytoma), reticulum cell sarcoma, Hodgkin's disease and
 CC tumours (e.g. glioma). The present sequence is one of the 1046
 CC novel cancer cell markers.
 XX
 XX Sequence 1245 BP; 382 A; 308 C; 390 G; 153 T; 12 other;
 SQ
 Query Match 26.8%; Score 179.6; DB 22; Length 1245;
 Best Local Similarity 67.1%; Pred. No. 6.1e-44;
 Matches 286; Conservative 0; Mismatches 134; Indels 6; Gaps 2;
 QY 217 GTGAATATGAGTTCGGCAGGAAGATCAACATATAGGCTAGGCCAGAGAAGATTACA 276
 Db 255 GTGAATATGAGTTCGGCAGGAGATCGAATATGGCTAGACCAGACGCTACGTAGA 314
 QY 277 GCCTCTGAGCTGATGGGCTATGCTTATCGGCTCCCTTTGTCAGGAACCCACTGA 336
 Db 315 GCCTCTGAAATGATTGGGCTTATGGGCC---CGAGCAGTTCAGTGATGAAGTGAACC 371
 QY 337 TGAAGAGCTTAAGAGAGAACCCACCACTAAAAGTCGGAATCCTACACCTGATCAGAA 396
 Db 372 AGCAACACCTGAAGAGGGGACCAAGCACTCAAGCTCAGGATCTGAGCTGCTCAGGA 431
 QY 397 GAGAGAAGATGATCAGGTCAGCTGAGATTCAGTGCCTGACCTGGAGCGGATCTCCA 456
 Db 432 GGGAGAGGATG---AGGAGAGCATCTGCAGGTCAAGGGCCGAGCCTGAAGCTCATAGCA 488
 QY 457 GAGGCTATCTCAGACAAAGACTGGGATGATGTGAAGTGTGATCTGATGTCAGGGGAA 516
 Db 489 GGAACAGGTCACCCACAGCTGGGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 548
 QY 517 GATTCTACAAAGCAGACACTTTTAAATGCGAGAGCAGGTGAAGGGAATCAAGGT 576
 Db 549 GGACCCGCCAATCCAGAGGAGGTGAAAACGGCTGAAGAAGGTGAAAAGCAATCAAGTG 608
 QY 577 TTAAGGAGATAGCTGAACACACAACTGTTTTATATAGATATTTTACTTTAA 636
 Db 609 TTAAGAAGAGGCACGTTGAAATGATGACGGCTCTCTATGTTGAAATTTGTTCAATTA 668
 QY 637 GAGTCT 642
 Db 669 AATTCT 674

RESULT 43
 AAF68852
 ID AAF68852 standard; cDNA; 461 BP.
 XX
 AC AAF68852;
 XX
 DT 12-APR-2001 (first entry)
 XX
 DE Human lung tumour protein related nucleotide sequence SEQ ID NO:792.
 XX
 KW Human; lung cancer; lung tumour; lung tumour protein; gene therapy;
 KW lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;

cytostatic; antisense inhibition; ss.

Homo sapiens.

WO200100828-A2.

04-JAN-2001.

30-JUN-2000; 2000WO-US18061.

30-JUN-1999; 99US-0346492.

15-OCT-1999; 99US-0419356.

17-DEC-1999; 99US-0468867.

30-DEC-1999; 99US-0476300.

06-MAR-2000; 2000US-0519642.

22-MAR-2000; 2000US-0533077.

10-APR-2000; 2000US-0546259.

27-APR-2000; 2000US-0560406.

05-JUN-2000; 2000US-0589184.

(CORI-) CORIXA CORP.

Wang T, Bangur CS, Lodes MJ, Fanger GR, Vedvick TS, Carter D;
 Retter MW, Mannion J;

WPI; 2001-071488/08.

Lung tumor-associated proteins and the nucleic acids that encode them,
 useful for preventing, diagnosing and treating lung cancer -

Claim 4; Page 413; 436pp; English.

The present invention describes immunogenic portions of lung tumour-
 associated proteins (I) and the nucleic acids (NAs) that encode them.
 (I) have cytostatic activity and can be used in gene therapy, antisense
 inhibition and in vaccines. The NAs and the lung tumour-associated
 proteins they encode may be used in the prevention, treatment and
 diagnosis of diseases associated with their inappropriate expression,
 especially lung cancers. For example, the NAs may be administered to
 treat diseases by rectifying mutations or deletions in a patient's genome
 that affect the activity of the protein by expressing inactive proteins
 or to supplement the patient's own production of (I). Additionally, the
 NAs may be used to produce the lung-tumour associated protein, according
 to standard recombinant DNA methodology. Conversely, antisense NA
 molecules may be administered to down regulate protein expression by
 binding with the cells own genes and preventing their expression. The NA
 and complementary sequences may also be used as DNA probes in diagnostic
 assays to detect and quantitate the presence of similar NA sequences in
 samples, and hence which patients may be in need of treatment for lung
 cancer. The (I) may be used as antigens in the production of antibodies
 and in assays to identify modulators (agonists and antagonists) of the
 expression and activity of the protein. AAF68083 to AAF68878 and
 AAB76948 to AAB76978 represent human lung tumour protein related
 nucleotide and protein sequences which are used in the exemplification
 of the present invention.

Sequence 461 BP; 150 A; 104 C; 123 G; 84 T; 0 other;

Query Match 26.8%; Score 178; DB 22; Length 461;

Best Local Similarity 73.9%; Pred. No. 1.1e-43;

Matches 260; Conservative 0; Mismatches 75; Indels 17; Gaps 2;

QY 319 GTCCAGGACCCACTGATCAAGAGCCTAAAGAGAGAGAACCCACCCTAAAGAGTGGAA 378
 Db 101 GTCCAGGACCCAGTATGAGAGCCGCCAAAAGAGAGACAGAGCTGAAGTGGGA 160
 QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCAGCTGAGATTC----- 428
 Db 161 TCCTACACCTGGGCAGCAGACAGAGAAGATCAGGATACAGCTCAGCTCAGGTGTGG 220
 QY 429 -----AAGTGCCTGACCTGGAGCCGATCTCCAGGAGCTATGTACACAAAGACTGGGG 482
 Db 221 GAAGGGAATGCGCGCATGGAAGTGTATTGGAAGAGCTGCATCAGTCAACACCGGG 280

QY 483 ATGATGTGAAGGT-GGTACTGATGTCAGGGGAAGATTCTACCAAAAGCAGACACTTT 541
 DB |||||
 QY 281 ATRAACTGCTGTTGGGTCGGCGTCAAGGTGAAGATAATACCTAAAGAGGAACACTGT 340
 DB |||||
 QY 542 AAAATGCCAAGCAGGTGAAGGGAATCAACAGTTTAAAGGAAGATAGCTGAACAC 601
 DB |||||
 QY 341 AAAATGCCAAGCAGGTGAAGGGAATCAACAGTTTAAAGGAAGATAGCTGAACAC 400
 DB |||||
 QY 602 ACAAACTGTTTTATATTAGATATTTTAAAGAGTCTTAATAAATTTT 653
 DB |||||
 QY 401 GCAAGCTGGTTTTATATTAGATATTTTAAAGAGTCTTAATAAATTTT 452
 DB |||||
 RESULT 44
 ID ABK38763 standard; cdNA; 461 BP.
 AC AC
 AC AC
 AC AC
 DT 21-MAY-2002 (first entry)
 XX
 XX cdNA encoding L552S lung tumour protein isoform.
 XX Lung tumour; cancer; T cell; immune response stimulator;
 XX cytostatic; gene; ss.
 XX Homo sapiens.
 OS
 XX WO200204514-A2.
 FN
 XX 17-JAN-2002.
 PD
 XX 10-JUL-2001; 2001WO-US22058.
 PF
 XX 11-JUL-2000; 2000US-0614124.
 PR 29-AUG-2000; 2000US-0651563.
 PR 08-SEP-2000; 2000US-0658824.
 PR 26-SEP-2000; 2000US-0671325.
 PR 06-OCT-2000; 2000US-0677419.
 PR 30-OCT-2000; 2000US-0702705.
 PR 13-DEC-2000; 2000US-0736457.
 PR 03-MAY-2001; 2001US-0849626.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 XX Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;
 PI Marxakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS;
 PI Mcnabb A, Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;
 XX
 DR WPI; 2002-164634/21.
 DR P-PSDB; AAU85524.
 XX
 PT Novel polynucleotide encoding a lung tumour polypeptide useful for
 PT stimulating and/or expanding T cells specific for a tumour protein -
 XX
 PS Example 1; SEQ ID No 792; 223pp; English.
 CC
 CC The invention describes an isolated polynucleotide and polypeptide
 CC useful for stimulating and/or expanding T cells specific for a tumour
 CC protein for determining the presence of a cancer in a patient. A
 CC composition containing the polynucleotide and/or polypeptide is useful
 CC for treating a lung cancer in a patient. The polypeptide is useful for
 CC removing tumour cells from a biological sample. The polynucleotide is
 CC also useful as probe or primer to detect the level of mRNA encoding a
 CC tumour protein. This sequence encodes a lung tumour associated protein
 CC or protein fragment, described in the method of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 461 BP; 150 A; 104 C; 123 G; 84 T; 0 other;

Query Match 26.6%; Score 178; DB 24; Length 461;
 Best Local Similarity 73.9%; Pred. No. 1.1e-43;
 Matches 260; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
 QY 319 GTCCAGGAACCCCTGATGAAGAGCCCTAAGAGAGAGAAACCCACTAAAGTCGGAA 378
 DB |||||
 QY 101 GTCCAGGAGCCAGTAAATGGAGAGCCCCCAAAAGAGAACAGCAGCTGAAGTCGGGA 160
 DB |||||
 QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGACGTGAGATTC----- 428
 DB |||||
 QY 161 TCCTACACCTGGCAGCAGACAGAGAGATCAGATACAGCTGAGATCCCGAGTGTGG 220
 DB |||||
 QY 429 -----AAGTGCTGACCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAGACTGGG 482
 DB |||||
 QY 221 GAAGGGAATGCGACATGGAAGTGATCTGCAAGAGCTGCATCAGTCNAACCCGGG 280
 DB |||||
 QY 483 ATGATGTGAAGGT-GGTACTGATGTCAGGGGAAGATTTTACCAAAAGCAGACACTTT 541
 DB |||||
 QY 281 ATRAACTGCTGTTGGGTTCCGGCTCAAGGTGAAGATAATACCTAAAGAGGAACACTGT 340
 DB |||||
 QY 542 AAAATGCCAAGCAGGTGAAGGGAATCAACAGTTTAAAGGAAGATAGCTGAACAC 601
 DB |||||
 QY 341 AAAATGCCAAGCAGGTGAAGGGAATCAACAGTTTAAAGGAAGATAGCTGAACAC 400
 DB |||||
 QY 602 ACAAACTGTTTTATATTAGATATTTTAAAGAGTCTTAATAAATTTT 653
 DB |||||
 QY 401 GCAAGCTGGTTTTATATTAGATATTTTAAAGAGTCTTAATAAATTTT 452
 DB |||||
 RESULT 45
 ID ACAL1092
 ID ACAL1092 standard; cdNA; 461 BP.
 AC ACAL1092;
 XX
 XX 05-JUN-2003 (first entry)
 DT
 DE Human lung cancer-associated cdNA L552S, isoform sequence.
 XX
 XX Human; lung cancer; ss; lung tumour; cytostatic; vaccine;
 KW T cell expansion; CD4; CD8.
 XX
 OS Homo sapiens.
 XX
 XX US2002197669-A1.
 FN
 XX 26-DEC-2002.
 PD
 XX 03-MAY-2001; 2001US-0849626.
 PF
 XX 13-DEC-2000; 2000US-0736457.
 PR
 XX (BANG/) BANGUR C S.
 PA (FANG/) FANGER G R.
 PA (WANG/) WANG A.
 PA (WANG/) WANG T.
 PA (SWIT/) SWITZER A P.
 PA (MCNE/) MCNEILL P D.
 PA (CLAP/) CLAPPER J D.
 XX
 XX Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;
 PI Clapper JD;
 XX
 DR WPI; 2003-352750/33.
 DR P-PSDB; ABU69497.
 XX
 PT Novel lung cancer polynucleotide encoding lung cancer protein, useful
 PT for detecting the presence of lung cancer in a patient, and in
 PT pharmaceutical compositions, e.g. vaccines, for treating lung cancer
 XX
 XX Example 1; Page -; 72pp; English.

Search completed: January 18, 2004, 01:58:21
Job time : 130.524 secs

The invention relates to a polynucleotide encoding a lung tumour protein, comprising a sequence selected from any of the 14 sequences mentioned in the specification, or a sequence (S2) mentioned in specification, complement of S1, sequences comprising of at least 20 contiguous residues of S1, sequences that hybridise to S1, sequences having 75%, preferably 90%, identity to S1, or degenerate variants of S1. Also included are an isolated polypeptide (comprising a sequence (S3) selected from any one of the 4 amino acid sequences mentioned in the specification, a sequence encoded by the polynucleotide, or sequences having at least 70%, preferably 90%, identity to a sequence encoded by the polynucleotide), an expression vector comprising the polynucleotide operably linked to an expression control sequence, a host cell transformed or transfected with the vector, an isolated antibody (or its antigen-binding fragment) that specifically binds to the polypeptide, detecting the presence of a cancer in a patient, a fusion protein comprising the polypeptide, an oligonucleotide that hybridises to S1 under moderately stringent conditions, stimulating and/or expanding T cells specific for a tumour protein (comprising contacting T cells with the polynucleotide, protein or antigen-presenting cells, under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells) and inhibiting the development of a cancer in a patient (by incubating CD4⁺ and/or CD8⁺ T cells isolated from a patient with the polynucleotide, protein or antigen presenting cells that express the polynucleotide, such that T cells proliferate, administering to the patient an effective amount of the proliferated T cells, and thus inhibiting the development of a cancer in the patient. The polynucleotide, protein and cells are useful in a composition for stimulating an immune response in a patient, and for treating a cancer in a patient (particularly lung cancer). The oligonucleotide is useful for determining the presence of a cancer in a patient. The protein and oligonucleotides are useful in pharmaceutical compositions, e.g. vaccines. The polynucleotide is also useful as a probe or primer for nucleic acid hybridisation, and in the design and preparation of ribozyme molecules for inhibiting expression of tumour polypeptides and proteins in tumour cells. An amplified portion of the polynucleotide is useful for isolating a full-length gene from a suitable library. The present sequence is a cDNA (full length, extended or partial) isolated from a library derived from lung tumour/cancer cells. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the USPTO <http://seqdata.uspto.gov/sequence.html?docid=20020197669>.

Query Match	26.6%	Score 178;	DB 25;	Length 461;
Best Local Similarity	73.9%;	Pred. No. 1.1e-43;		
Matches	260;	Conservative 0;	Mismatches 75;	Indels 17; Gaps 2;
Qy	319	GTCCAGGACCCACTGATGAGAGGCTTAAGAGAGGAAACACACCACCTAAAGATCGGAA	378	
Db	101	GTCCAGGAGCCCGATTAATGAGAGCCCCAAAAGAGAAACCGACAGCTGAAGTCGGGA	160	
Qy	379	TCTTACACCTGATCAGAAGAGAGATGATCAGGCTGCAGCTGAGATTCTC-----	428	
Db	161	TCTTACACCTGGGCGAGCGACAGAGAGAGATCAGGATACAGCTAGATCCAGGTGCTGG	220	
Qy	429	-----AAGTGCTGACTGTGAAGCCGATCTCCAGGAGCTATGTCTAGAATAAGACTGGGG	482	
Db	221	GAAGGGAATGCGCACATGGAGGTGATCTGCAAGAGCTGCACTCAGTCAAAACACCGGGG	280	
Qy	483	ATGGATGTGAAGGT-GGTACTGATGTCAAGGGGAGATTCTACCAAAACGACAGACTTT	541	
Db	281	ATAAATCTGGATTGGTTTCGGCGCTCAAGGTGAAGATAATACCTTAAGAGGACACTGT	340	
Qy	542	AAAATGCCAGACAGGTGAAGGGAAATCAAGGTTTTAAAGGAAGATAGCTTGAACACAC	601	
Db	341	AAATGCCAGACAGGTGAAGAGCAACCAAGTTTTAATGAAGACAGCTGAACACAC	400	
Qy	602	ACAAACTGTTTTTATATTAGATATTTTTACTTTAAAGAGTCTTTAATAATTTT	653	
Db	401	GCAAGCTGGTTTTTATATTAGATATTGTTGACTTTAAACTATCTCAATAAAGTTTT	452	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 18, 2004, 01:12:25 ; Search time 32.4394 Seconds
(without alignments)
9102.683 Million cell updates/sec

Title: US-10-051-835-1

Perfect score: 669
Sequence: 1 cacaacgcagccagcaatt.....tttttggcatgctgatctc 669

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139556

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTCUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	317	47.4	659	4	US-08-620-312D-634
2	204	30.5	399	4	US-08-702-705-69
3	204	30.5	399	4	US-08-736-457-69
4	204	30.5	457	4	US-08-702-705-790
5	204	30.5	457	4	US-08-736-457-790
6	204	30.5	781	4	US-08-702-705-808
7	204	30.5	781	4	US-08-736-457-808
8	182.8	27.3	528	4	US-08-163-748C-1
9	181.2	27.1	538	1	US-08-370-648-14
10	181.2	27.1	538	2	US-08-531-662B-14
11	181.2	27.1	538	3	US-08-602-039-14
12	181.2	27.1	538	3	US-08-602-039-14
13	180	26.9	479	4	US-08-702-705-70
14	180	26.9	479	4	US-08-736-457-70
15	178	26.6	461	4	US-08-702-705-792
16	178	26.6	461	4	US-08-736-457-792
17	169.8	25.4	532	1	US-08-370-648-17
18	169.8	25.4	532	2	US-08-531-662B-17
19	169.8	25.4	532	3	US-08-669-161A-17
20	169.8	25.4	532	3	US-08-602-039-17
21	169.8	25.4	539	1	US-08-370-648-18
22	169.8	25.4	539	2	US-08-531-662B-18
23	169.8	25.4	539	3	US-08-669-161A-18
24	169.8	25.4	539	3	US-08-602-039-18
25	169.8	25.4	540	1	US-08-370-648-16
26	169.8	25.4	540	2	US-08-531-662B-16
27	169.8	25.4	540	3	US-08-669-161A-16

28	169.8	25.4	540	3	US-08-602-039-16
29	168.2	25.1	526	4	US-09-163-748C-2
30	162.2	24.2	560	1	US-08-370-648-15
31	162.2	24.2	560	2	US-08-531-662B-15
32	162.2	24.2	560	3	US-08-669-161A-15
33	162.2	24.2	560	3	US-08-602-039-15
34	155.2	23.2	288	4	US-09-702-705-1827
35	155.2	23.2	288	4	US-09-736-457-1827
36	148.6	22.2	656	4	US-09-328-475C-263
37	144.2	21.6	646	1	US-08-370-648-1
38	144.2	21.6	646	2	US-08-531-662B-1
39	144.2	21.6	646	3	US-08-669-161A-1
40	144.2	21.6	646	3	US-08-602-039-1
41	144.2	21.6	648	1	US-08-250-162A-1
42	142.8	21.3	1024	4	US-09-328-475C-78
43	140.8	21.0	1024	4	US-09-328-475C-77
44	99.2	14.8	515	4	US-08-439-313-472
45	99.2	14.8	515	4	US-09-352-616A-472
46	97	14.5	9531	4	US-09-163-748C-3
47	97	14.5	11461	3	US-08-669-161A-29
48	78	11.7	1092	4	US-09-712-016-5
49	76.2	11.4	141	4	US-09-702-705-1828
50	76.2	11.4	141	4	US-09-736-457-1828
51	60.6	9.1	7218	1	US-08-232-463-14
52	42.8	6.4	111	4	US-09-702-705-1829
53	42.8	6.4	111	4	US-09-736-457-1829
54	39.6	5.9	289	3	US-09-007-005-17
55	39.6	5.9	289	3	US-08-766-738-2
56	35.4	5.3	966	2	US-08-244-796-17
57	35.4	5.3	966	2	US-09-262-610-2
58	34.6	5.2	246	4	US-09-407-956-3
59	34.6	5.2	888	4	US-09-407-956-1
60	33.6	5.0	28473	4	US-08-961-527-83
61	33.4	5.0	13608	4	US-09-539-333D-1
62	33.4	5.0	123025	4	US-09-198-452A-1
63	33.2	5.0	6678	4	US-09-268-347-33
64	33	4.9	822	3	US-08-896-164-47
65	33	4.9	1344	4	US-09-480-921B-5
66	33	4.9	4140	4	US-09-480-921B-19
67	33	4.9	7739	3	US-09-195-966-1
68	33	4.9	7745	1	US-08-299-675-1
69	33	4.9	7745	1	US-08-485-241-1
70	33	4.9	7745	2	US-08-874-162-1
71	33	4.9	61663	4	US-09-453-702B-62
72	32.8	4.9	325	1	US-08-650-275-8
73	32.8	4.9	325	3	US-09-181-318-8
74	32.8	4.9	374	1	US-08-341-538A-3
75	32.8	4.9	374	2	US-08-725-518-3
76	32.8	4.9	1113	1	US-08-341-538A-1
77	32.8	4.9	1113	2	US-08-725-518-1
78	32.8	4.9	1323	4	US-09-252-991A-5119
79	32.8	4.9	1422	4	US-09-252-991A-5119
80	32.8	4.9	1467	4	US-09-252-991A-5070
81	32.8	4.9	1664976	4	US-08-916-421B-1
82	32.6	4.9	5361	3	US-08-973-462-2
83	32.6	4.9	5361	3	US-08-973-462-1
84	32.2	4.8	277	4	US-09-313-294A-6369
85	32.2	4.8	319608	4	US-09-679-409-1
86	32	4.8	1551	1	US-08-485-241-4
87	32	4.8	1551	2	US-08-874-162-4
88	32	4.8	1919	1	US-08-485-241-2
89	32	4.8	1919	2	US-08-874-162-2
90	32	4.8	3334	1	US-09-996-243-288
91	31.8	4.8	394	1	US-08-650-275-12
92	31.8	4.8	394	3	US-09-181-318-12
93	31.8	4.8	2277	1	US-08-676-967-2
94	31.8	4.8	2277	1	US-08-676-974-2
95	31.8	4.8	2277	2	US-09-096-487-2
96	31.8	4.8	3592	3	US-08-714-918-63
97	31.8	4.8	3592	3	US-09-265-315-63
98	31.8	4.8	3592	3	US-09-265-315-63
99	31.8	4.8	3592	3	US-09-266-417-63
100	31.6	4.7	429	4	US-09-643-597-199

Sequence 16, Appl
Sequence 2, Appl
Sequence 15, Appl
Sequence 15, Appl
Sequence 15, Appl
Sequence 15, Appl
Sequence 1827, Ap
Sequence 1827, Ap
Sequence 263, App
Sequence 1, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 77, Appl
Sequence 78, Appl
Sequence 472, App
Sequence 472, App
Sequence 3, Appl
Sequence 29, Appl
Sequence 5, Appl
Sequence 1828, Ap
Sequence 1828, Ap
Sequence 14, Appl
Sequence 1829, Ap
Sequence 1829, Ap
Sequence 17, Appl
Sequence 2, Appl
Sequence 3, Appl
Sequence 83, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 62, Appl
Sequence 8, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 5147, Ap
Sequence 5119, Ap
Sequence 5070, Ap
Sequence 1, Appl
Sequence 2, Appl
Sequence 1, Appl
Sequence 6369, Ap
Sequence 1, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 288, App
Sequence 12, Appl
Sequence 12, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 63, Appl
Sequence 63, Appl
Sequence 63, Appl
Sequence 199, App

ALIGNMENTS

RESULT 1

US-03-620-312D-634
; Sequence 634, Application US/09620312D
; Patent No. 6569662

GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; POLYPEPTIDES
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt FL_Genes Version 1.0
; SEQ ID NO 634
; LENGTH: 659
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (225)..(560)
US-09-620-312D-634

Query Match 47.4%; Score 317; DB 4; Length 659;

Best Local Similarity 77.4%; Pred. No. 1.6e-91;

Matches 442; Conservative 0; Mismatches 85; Indels 44; Gaps 3;

```
QY 86 CTTTGGCCACGTGTGACCGCTGGGAGCTGTGAGAGTGTGAGGGGACGCTTCCAGCG 145
   |||
Db 109 CTTTTCACCGTGTGATGCTGAGCTGTGAGGGTGTGAGGGTGTGAGGGTGTGAGGGT 168

QY 146 TCTGGACTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 205
   |||
Db 169 TCTGACATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 210

QY 206 GAACCTGAATAGTGAATATAGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 265
   |||
Db 211 ----TTGCAGTGTGAATATAGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 266

QY 266 AGAAGTTACAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 325
   |||
Db 267 AGAAGTTACAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 305

QY 326 GAACCCACTGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 385
   |||
Db 306 GAGCCCGGTGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 365

QY 386 CCGTGTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 445
   |||
```

RESULT 2

US-09-702-705-69

; Sequence 69, Application US/09702705

; Patent No. 6504010

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Lodes, Michael A.

; APPLICANT: Panger, Gary

; APPLICANT: Vedwick, Tom

; APPLICANT: Carter, Darrick

; APPLICANT: Retter, Marc

; APPLICANT: Mannion, Jane

; APPLICANT: Fan, Liqun

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.478C14

; CURRENT APPLICATION NUMBER: US/09/702,705

; CURRENT FILING DATE: 2000-10-30

; NUMBER OF SEQ ID NOS: 1833

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 69

; LENGTH: 399

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-702-705-69

Query Match 30.5%; Score 204; DB 4; Length 399;

Best Local Similarity 77.4%; Pred. No. 1.7e-55;

Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

```
QY 319 GTCCGAGGAGCCCACTGATGAAGAGGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAG 378
   |||
Db 40 GTCCGAGGAGCCCACTGATGAAGAGGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAG 99

QY 379 TCCTACACCTGATCAGAAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGA 438
   |||
Db 100 TCCTACACCTGGCAGCAGACAGAGAGATCAGATACAGCTGAGATCCCACTGGCGA 159

QY 439 CTGGAAGCCGATCTCCAGGAGCTATGTGAGACAAAGATGCGGGATGGATGTGAGAGGT-G 497
   |||
Db 160 CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCCGGGATAAATCTGGATTGG 219

QY 498 GTACTCATGTCAAGGGGAAGATTCTTACAAAAGCAGAGACACTTTAAATGCCAGAGCAG 557
   |||
Db 220 GTTCCGGCTCAAGGTGAAGATAATACCTTAAGAGGAGACACTGTAAATGCCAGAGCAG 279

QY 558 GTGAAGGGAAATCAGAGGTTTAAAGGAAGATTAAGCTGAAGAACCAACAACTGTTTTATA 617
   |||
Db 280 GTGAAGAGCAACCACTAAGTTTAAATGAAGACAGCTGAAGAACCAACCACTGTTTTATA 339

QY 618 TTAGATATTTTACTTTTAAAGAGTCTTTAATAAATTTT 653
   |||
Db 340 TTAGATATTGACTTAAACTATCTCAATAAAGTTTTT 375
```


RESULT 3

US-09-736-457-69
; Sequence 69, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 69
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-69

Query Match 30.5%; Score 204; DB 4; Length 399;
Best Local Similarity 77.4%; Pred. No. 1.7e-55;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

QY 319 GTCCAGGAACCCACTGATGAAGAGCCTAAAGAGAGAACCCACCCACTAAAGAGTCGGAA 378
DB 40 GTCCAGGAGCCAGTAAATGGAGAGGCCCAAAAGAGAACCCAGCAGCTGAAGTCGGGA 99

QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGCGAGCTCAGATTCAAGTCGCTGA 438
DB 100 TCCTACACCTGGGAGAGAGAGATGATCAGAGAGAGATCAGGATACAGTCCAGTCGGCGA 159

QY 439 CCTGGAAGCCGATCTCCAGAGAGCTATGTCAGACAAAGACTGGGGATGGATGTGAAGT-G 497
DB 160 CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAACACCCGGGATAAATCTCGGATTTGG 219

QY 498 GTACTGATGTCAAGGGAAGATTCTACAAAGCAGAGCAGCTTTAAATGCCAGAGCAG 557
DB 220 GTTCGGGTCAGGTGAAGATATATCTAAAGAGGAACTGTAAATGCCAGAGCAG 279

QY 558 GTGAAGGGAATACACAGTTTAAAGGAGATAAGCTGAAACAAACACAACTGTTTTTATA 617
DB 280 GTGAAGAGCAACCAAGTTTAAATGAAGACAGCTGAAACAAAGCTGTTGTTTTATA 339

QY 618 TTAGTATTTTACTTTAAAGAGTCTTAATAATTTT 653
DB 340 TTAGTATTTGACTTAAACTATCTCAATAAAGTTTT 375

RESULT 4

US-09-702-705-790
; Sequence 790, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 790
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-790

Query Match 30.5%; Score 204; DB 4; Length 457;
Best Local Similarity 77.4%; Pred. No. 1.8e-55;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

QY 319 GTCCAGGAACCCACTGATGAAGAGCCTAAAGAGAGAACCCACCCACTAAAGTCGGAA 378
DB 120 GTCCAGGAGCCAGTAAATGGAGAGGCCCAAAAGAGAACCCAGCAGCTGAAAGTCGGGA 179

QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGCGAGCTGAGATTCAAGTCGCTGA 438
DB 180 TCCTACACCTGGGAGAGAGAGATGATCAGGATACAGTCCAGTCGCGGA 239

QY 439 CCTGGAAGCCGATCTCCAGAGAGCTATGTCAGACAAAGACTGGGGATGGATGTGAAGT-G 497
DB 240 CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAACACCCGGGATAAATCTGGATTTGG 299

QY 498 GTACTGATGTCAAGGGAAGATTCTACAAAGCAGAGCAGCTTTAAATGCCAGAGCAG 557
DB 300 GTTCGGGTCAGGTGAAGATATATCTAAAGAGGAACTGTAAATGCCAGAGCAG 359

QY 558 GTGAAGGGAATACACAGTTTAAAGGAGATAAGCTGAAACAAACACAACTGTTTTTATA 617
DB 360 GTGAAGAGCAACCAAGTTTAAATGAAGACAGCTGAACAAACGCAAGCTGGTTTTATA 419

QY 618 TTAGTATTTTACTTTAAAGAGTCTTAATAATTTT 653
DB 420 TTAGTATTTGACTTAAACTATCTCAATAAAGTTTT 455

RESULT 5

US-09-736-457-790
; Sequence 790, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 790
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-736-457-790

Query Match 30.5%; Score 204; DB 4; Length 457;
Best Local Similarity 77.4%; Pred. No. 1.8e-55;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;


```
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Which Encode Member Of The Gage
; FILE REFERENCE: LUD 5558
; CURRENT APPLICATION NUMBER: US/09/163,748C
; CURRENT FILING DATE: 1998-09-30
; NUMBER OF SEQ ID NOS: 27
; SEQ ID NO 1
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-163-748C-1

Query Match      27.3%; Score 182.8; DB 4; Length 528;
Best Local Similarity 67.4%; Pred. No. 1.2e-48;
Matches 289; Conservative 0; Mismatches 132; Indels 6; Gaps 2;

QY 217 GTGAATATGAGTTGGCGAGAGATCAACATATAGCCCTAGCCCAAGAGAGTTTACA 276
Db 66 GTGAATATGAGTTGGCGAGAGATCGACCTATCGGCTAGACCAGAGCTACGTAGA 125

QY 277 GCTCTCTGAGCTGATTGGGCTATGTTACTTGCTCCCTTTGTCCTCCAGAACCCACTGA 336
Db 126 GCCTCTCTGAAATGATTGGGCTATGCGGCC---CGAGCAGTTTCAGTGAATGGAACC 182

QY 337 TGAAGAGCCTAAAGAGAGAGAAACCCACCCTAAAGTCGGAATCCTACACCTGATCAGAA 396
Db 183 AGCAACACCTGAAGAGAGAGAACAGCACTCAAGTCAGGATCTCTGAGCTGCTCAGA 242

QY 397 GAGAGAGATGATCAGGGTCCAGCTGAGATTCAGTTCAGTCCCTGAGCTCGAAGCCGATCTCA 456
Db 243 GGGAGAGAGATG---AGGAGCAGCTTCGAGGTCGAAGGCCCGAAGCCTGAACTGATAGCCA 299

QY 457 GGAGCTATGTCAGCAAGAGCTGGGATGGATGATGAGTGGTGTGAGTGGTGTGAGTGGGAGAA 516
Db 300 GGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAGTGGTGTGAGTGGGAGAGAT 359

QY 517 GATTCTACCAAGAGCAGACACTTTAAATGCCAGAGCAGGTGGAAGGGAATCAGAGT 576
Db 360 GGACCCGCCAAATCCAGAGAGGTGAAACGCCCTGAAGAGGTGAAAGCAATCAGAGT 419

QY 577 TTAAGAGAGATAGCTGAACACACAACTGTTTATATAGATATTTTACTTTAA 536
Db 420 TTAAGAGAGACAGCTTGAATGATGACAGGCTGCTCTATGTTGGAATTTGTTCAATTA 479

QY 637 GAGTCT 642
Db 480 AATTCT 485

RESULT 9
US-08-370-648-14
; Sequence 14, Application US/08370648
; Patent No. 5648226
; GENERAL INFORMATION:
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry;
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
```

```
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,648
; FILING DATE: 10-JANUARY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-JULY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5648226man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5323.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 538 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-370-648-14

Query Match      27.1%; Score 181.2; DB 1; Length 538;
Best Local Similarity 67.4%; Pred. No. 4e-48;
Matches 287; Conservative 0; Mismatches 133; Indels 6; Gaps 2;

QY 217 GTGAATATGAGTTGGCGAGAGATCAACATATAGCCCTAGCCCAAGAGAGTTTACA 276
Db 77 GTGAATATGAGTTGGCGAGAGATCGACCTATCGGCTAGACCAAGACGCTACGTAGA 136

QY 277 GCCTCTCTGAGCTGATTGGGCTATGCTTACTTGCTCCCTTTGTCCTCCAGAACCCACTGA 336
Db 137 GCCTCTCTGAAATGATTGGGCTATGCGGCC---CGAGCAGTTTCAGTGAATGGAACC 193

QY 337 TGAAGAGCCTAAAGAGAGAGAAACCCACCCTAAAGTCGGAATCCTACACCTGATCAGAA 396
Db 194 AGCAACACCTGAAGAGAGAGAAACCCAGCAACTCAACAGCTCAGGATCCTGCGAGCTGCTCAGA 253

QY 397 GAGAGAGATGATCAGGGTGCAGCTGAGATTCAGTTCAGTCCCTGAGCCTGGAAGCCGATCTCCA 456
Db 254 GGGAGAGAGATG---AGGAGCAGTCTGAGGTCAAGGCCCGAGCCTGAGCTCATAGCCA 310

QY 457 GGAGCTATGTCAGACAAAGACTGGGGATGGATGTAAGGTGGTGTGAGTGGTGTGAGTGGGAGAGAT 516
Db 311 GGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAGTGGTGTGAGTGGGAGAGAT 370

QY 517 GATTCTACCAAGAGCAGACACTTTAAATGCCAGAGCAGGTGGAAGGGAATCAGAGT 576
Db 371 GGACCCGCCAAATCCAGAGAGGTGAAACGCCCTGAAGAGAGGTGAAAGCAATCAGAGT 430

QY 577 TTAAGAGAGATAGCTGAAACACACAACTGTTTTATATAGATATTTTACTTTAA 636
Db 431 TTAAGAGAGACAGCTTGAATGATGACAGGCTGCTCTATGTTGGAATTTGTTCAATTA 490

QY 637 GAGTCT 642
Db 491 AATTCT 496

RESULT 10
US-08-531-662B-14
; Sequence 14, Application US/08531662B
; Patent No. 5858689
; GENERAL INFORMATION:
; APPLICANT: Van der Bruggen, Pierre;
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: Isolated, Truncated Nucleic
; TITLE OF INVENTION: Acid Molecules Which Code For GAGE Tumor
```

;; TITLE OF INVENTION: Rejection Antigen, The Tumor Rejection Antigen,
;; TITLE OF INVENTION: and Uses Thereof
;; NUMBER OF SEQUENCES: 23
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Felfe & Lynch
;; STREET: 805 Third Avenue
;; CITY: New York City
;; STATE: New York
;; COUNTRY: USA
;; ZIP: 10022
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
;; OPERATING SYSTEM: PC-DOS
;; SOFTWARE: Wordperfect
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/531.662B
;; FILING DATE: 21-September-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/370,648
;; FILING DATE: 10-January-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/250,162
;; FILING DATE: 27-May-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/096,039
;; FILING DATE: 22-July-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hanson, No. 585689man D.
;; REGISTRATION NUMBER: 30,946
;; REFERENCE/DOCKET NUMBER: LUD 5323.3
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 688-9200
;; TELEFAX: (212) 838-3884
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 538 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; US-08-531-662B-14
Query Match 27.1%; Score 181.2; DB 2; Length 538;
Best Local Similarity 67.4%; Pred. No. 4e-48;
Matches 287; Conservative 0; Mismatches 133; Indels 6; Gaps 2;
QY 217 GTGAATATGAGTTGGCGAGGAAGATCAACATATAGGCTAGGCGCCAGCAAGAGTTTACA 276
Db 77 GTGAATATGAGTTGGCGAGGAAGATCGACCTATCGGCTAGACCAAGCGCTAGTAGA 136
QY 277 GCCTCTGAGCTGATGGGGCTATGCTTACTGGCTCCCTTTGTCCCGAGAACCCACTGA 336
Db 137 GCCTCTGAAATGATTGGGCGCTATCGGCC---CGAGCAGTTCAGTGATGAAGTGAACC 193
QY 337 TGAAGAGCCTAAAGAAGAAACCAACCACTAAAAGTGGGAATCTACACCTGATCAGAA 396
Db 194 AGCAACACTGAAAGAGGGACCACTCAACGTCAGGATCTGCGAGTCTGCTCAGGA 253
QY 397 GAGAGAAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAGCGGATCTCCA 456
Db 254 GGGAGAGGATG---AGGAGCATCTGCGAGTCAAGGGCGGAGCGCTGAAGCTCATAGCA 310
QY 457 GGAGCTATCTCAGACAAAGACTGGGATGGAATGGAAGTGTGATGATGTCAGAGGGAA 516
Db 311 GGNACAGGTCACCCACAGCTGGGTGTGAGTGTGAGATGCTCTGATGGCAGGAT 370
QY 517 GATTTCACAAAGCAGACATTTAAATGCCAAGCAGGTGAAGGGAATCAAGGT 576
Db 371 GGACCGCCAAATCCAGAGGAGTGAACACGCGCTGAAGAGGTGAAAGCAATCAAGTG 430
QY 577 TTAAGGAGATGAGCTGAACACAACTGTTTATATATAGATATTTTACTTTTAA 636

Db 431 TTAAGAAGACACGTTGAATGATGAGGCTGCTCTATGTTGGAATTTGTTCAATAA 490
QY 637 GAGTCT 642
Db 491 AATTCT 496
RESULT 11
US-08-669-161A-14
; Sequence 14, Application US/08669161A
; Patent No. 6013481
; GENERAL INFORMATION:
; APPLICANT: DeBacker, Olivier; Van den Eynde,
; APPLICANT: Benoit; Boon-Palleur, Thierry
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/669,161A
; FILING DATE: 24-June-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/531,662
; FILING DATE: 21-September-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/370,648
; FILING DATE: 10-January-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-May-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-July-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6013481man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5443
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 538 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-669-161A-14
Query Match 27.1%; Score 181.2; DB 3; Length 538;
Best Local Similarity 67.4%; Pred. No. 4e-48;
Matches 287; Conservative 0; Mismatches 133; Indels 6; Gaps 2;
QY 217 GTGAATATGAGTTGGCGAGGAAGATCAACATATAGGCTAGGCGCCAGCAAGAGTTTACA 276
Db 77 GTGAATATGAGTTGGCGAGGAAGATCGACCTATCGGCTAGACCAAGCGCTAGTAGA 136
QY 277 GCCTCTGAGCTGATGGGGCTATGCTTACTGGCTCCCTTTGTCCCGAGAACCCACTGA 336
Db 137 GCCTCTGAAATGATTGGGCGCTATCGGCC---CGAGCAGTTCAGTGATGAAGTGAACC 193

Qy	379	TCCTACCTGNTCAGACAGAGAGAGATGATCAGGGTGCAGCTGAGATTC-----	428
Db	163	TCTTACCTCTGGCAGCAGACAGAAAGATCAGATACAGCTTGAGTCCAGGTCGTGG	222
Qy	429	-----AAGTGCTGACTGGAGCCGATCTCCAGGAGCTATGTCTCAGACAAAGACTGGG	482
Db	223	GAAAGGAAATGCGCAGCATGGAAGTGATCTGCAAGAGCTGCGATCAGTCAACACCGGG	282
Qy	483	ATGGATGTGAAGGT--GGTACTGATGTCAAGGGGAAGATCTTCACAAAGCAGAGCACTTT	541
Db	283	ATAAATCTCGATTTTGGGTTCGGCGGTCAAGGTGAAGATAATACCTAAAGAGGAACACTGT	342
Qy	542	AAAAATCCAGAAGCAGGTTGAAGGGAAATCACAGGTTTTAAAGGGAAGATAAGCTGAAACAC	601
Db	343	AAAATGCCAGAAGCAGGTTGAAGGCAACCAACAAGTTTTAAATGAAGACAAGCTGAAACAC	402
Qy	602	CAAACTGTTTTTATTATTAGATATTTTACTTTTAAAGAGTCTTAATAAAATTTTGG	655
Db	403	GCAAGCTGGTTTTTATTATTAGGATATTTGACTTTAAACTATCTCAATAAAGTTTGG	456

```

RESULT 14
US-09-736-457-70
; Sequence 70, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 70
; LENGTH: 479
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-70

```

```

Cy 602 ACAACGCTTTTATATAGATATTTTACTTTAAAGAGCTCTTAATAAATTTTG 655
Db 403 GCAAGCTGGTTTATATAGATATTTGACTTAAACTATCTCAATAAGTTTG 456

```

RESULT 16
US-09-736-457-792
; Sequence 792, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Barrick

STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/531,662B
FILING DATE: 21-September-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/370,648
FILING DATE: 10-January-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250,162
FILING DATE: 27-May-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/096,039
FILING DATE: 22-July-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 585868man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5323.3
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-531-662B-17

Query Match 25.4%; Score 169.8; DB 2; Length 532;
Best Local Similarity 67.1%; Pred. No. 1.8e-44;
Matches 288; Conservative 0; Mismatches 132; Indels 9; Gaps 3;

Qy 217 GTGAATATGATGGCGAGAGATCAAC--ATATAGCCTAGGCCAGGAGAGTTT 273
Db 68 GTGAATATGATGGCGAGAGATCGACCTATTATGGCCTAGACCAGGCGCTATGT 127

Qy 274 ACAGCCTCTGAGCTGATGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAACCCAC 333
Db 128 ACAGCCTCTGAGCTGATGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAACCCAC 184

Qy 334 TGATGAGAGCCTAAAGAGAGAACCCACCACTAAAGTCGGAATCTTACCTGATCA 393
Db 185 ACCAGCAACCTGAAAGAGAGAACCCACCACTAAAGTCGGAATCTTACCTGATCA 244

Qy 394 GAAGAGAGAGATGATCAGGCTGAGATTCAAGTGCCTGACCTGGAAGCCGATCT 453
Db 245 GGAGGAGAGATG--AGGAGCATCTGAGTCAAGGCCGAGCCTGAGCTGATAG 301

Qy 454 CCAGGAGCTATGTCAGCAAGAGTGGGATGATGTAAGTGGTACTGATGTCAGG 513
Db 302 CCAGGAGAGGCTCACCAGAGTGGGCTGATGTAAGTGGTCTGATGGCAGGA 361

Qy 514 GAAGATTTACCAAAAGCAGAGCACTTTAAATGCCAGAGAGGTCGAGGAAATCACA 573
Db 362 GATGACCCGCAATCCAGAGAGGTGAAACCGCTGAGAGAGGTGAAAGCAATCACA 421

Qy 574 GGTTTAAAGGAGATAGCTGAAACACACAACTGTTTTTATATAGATATTTTACTTT 633
Db 422 GTGTTAAAGGAGGACGCTTGAATGATGATGAGGCTGCTCTATGTTGAAATTTGTTCA 481

Qy 634 AAGAGTCT 642
Db 482 TAAATTTCT 490

RESULT 19
US-08-669-161A-17
Sequence 17, Application US/08669161A
Patent No. 6013481
GENERAL INFORMATION:
APPLICANT: Debacker, Olivier; Van den Eynde,
APPLICANT: Benoit; Boon-Falleur, Thierry
TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pelfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/669,161A
FILING DATE: 24-June-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/531,662
FILING DATE: 21-September-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/370,648
FILING DATE: 10-January-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250,162
FILING DATE: 27-May-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/096,039
FILING DATE: 22-July-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6013481man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5443
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-669-161A-17

Query Match 25.4%; Score 169.8; DB 3; Length 532;
Best Local Similarity 67.1%; Pred. No. 1.8e-44;
Matches 288; Conservative 0; Mismatches 132; Indels 9; Gaps 3;

Qy 217 GTGAATATGATGGCGAGAGATCAAC--ATATAGCCTAGGCCAGGAGAGTTT 273
Db 68 GTGAATATGATGGCGAGAGATCGACCTATTATGGCCTAGACCAGGCGCTATGT 127

Qy 274 ACAGCCTCTGAGCTGATGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAACCCAC 333
Db 128 ACAGCCTCTGAGCTGATGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAACCCAC 184

Qy 334 TGATGAGAGCCTAAAGAGAGAACCCACCACTAAAGTCGGAATCTTACCTGATCA 393
Db 185 ACCAGCAACCTGAAAGAGAGAACCCACCACTAAAGTCGGAATCTTACCTGATCA 244

Qy 394 GAAGAGAGAGATGATCAGGCTGAGATTCAAGTGCCTGACCTGGAAGCCGATCT 453
Db 245 GGAGGAGAGATG--AGGAGCATCTGAGTCAAGGCCGAGCCTGAGCTGATAG 301

Qy 454 CCAGGAGCTATGTCAGCAAGAGTGGGATGATGTAAGTGGTACTGATGTCAGG 513
Db 302 CCAGGAGAGGCTCACCAGAGTGGGCTGATGTAAGTGGTCTGATGGCAGGA 361

Qy 514 GAAGATTTACCAAAAGCAGAGCACTTTAAATGCCAGAGAGGTCGAGGAAATCACA 573
Db 362 GATGACCCGCAATCCAGAGAGGTGAAACCGCTGAGAGAGGTGAAAGCAATCACA 421

Qy 574 GGTTTAAAGGAGATAGCTGAAACACACAACTGTTTTTATATAGATATTTTACTTT 633
Db 422 GTGTTAAAGGAGGACGCTTGAATGATGATGAGGCTGCTCTATGTTGAAATTTGTTCA 481

Qy 634 AAGAGTCT 642
Db 482 TAAATTTCT 490

Db 245 GGAGGAGAGGATG---AGGAGCATCTGCAGGTCAAGGCCGAGACCTGAAAGCTGATAG 301
Qy 454 CCAGGAGCTATGTCAGACAAAGCTGGGATGGATGTGAAGTGTGACTGATGTCAAGG 513
Db 302 CCAGGAAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGTGTCTCTGATGGGCAGGA 361
Qy 514 GAAGATTCTACCAAAAGCAGACACTTTAAATGCCAGAGCAGGTGAAGGGAATATCACA 573
Db 362 GATGGACCCGCCAATCCAGAGAGGTGAACGCCCTGAAGAGGTGAAGCAATCACA 421
Qy 574 GGTTTAAAGGAGATAGCTGAAACACACAACTGTCTTTTATATATAGATATTTTACTTT 633
Db 422 GTGTTAAAGAGGACGTTGAATATGATGAGCTGCTCTCTATGTGGAAATTTGTTAT 481
Qy 634 AAGAGCTCT 642
Db 482 TAAATTTCT 490

RESULT 20
US-08-602-039-17
; Sequence 17, Application US/08602039
; Patent No. 6069001
; GENERAL INFORMATION:
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry;
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,039
; FILING DATE: 15-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/370,648
; FILING DATE: 10-JANUARY-1995
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-JULY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6069001man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5323.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-602-039-17

Query Match 25.4%; Score 169.8; DB 3; Length 532;
Best Local Similarity 67.1%; Pred. No. 1.8e-44;
Matches 288; Conservative 0; Mismatches 132; Indels 9; Gaps 3;

Qy 217 GTGAATATACAGTTCGGCAGGAATCAAC---ATATAGGCTAGGCCAAGAGAGAGTTT 273
Db 68 GTGAATATACAGTTCGGCAGGAAGATCGACCTATTATTGGCCCTAGACCAAGGCGCTAATGT 127
Qy 274 ACAGCCTCCTGAGCTGATTGGGGCTATGCTTTACTTGGCTCCCTTTTGTCCAGGAACCCAC 333
Db 128 ACAGCCTCCTGAGTGTATTGGGCTATGCGCC---CGAGCAGTTCAAGTATGAAGTGA 184
Qy 334 TGATGAAGAGCTTAAAGAGAGAAACCCACCTAAAGTTCGGAATCTACACTGATCA 393
Db 185 ACCAGCAACACCTGAAGAGGGAAACCCAGCAACTCAACGTCAGGATCCTGAGCTGCTCA 244
Qy 394 GAAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAGATTCAGCTGCCTGAAAGCCGATCT 453
Db 245 GGAGGGAGAGGATG---AGGAGCATCTGCAGGTCAAGGCCGCAAGCCTGAAGCTGATAG 301
Qy 454 CCAGGAGCTATGTCAGACAAAGACTGGGATGGATGTGAAGTGTGACTGATGTCAAGG 513
Db 302 CCAGGAAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGTGTCTCTGATGGGCAGGA 361
Qy 514 GAAGATTCTACCAAAAGCAGACACTTTAAATGCCAGAGCAGGTGAAGGGAATATCACA 573
Db 362 GATGGACCCGCCAATCCAGAGAGGTGAACGCCCTGAAGAGGTGAAGCAATCACA 421
Qy 574 GGTTTAAAGGAGATAGCTGAAACACACAACTGTCTTTTATATATAGATATTTTACTTT 633
Db 422 GTGTTAAAGAGGACGCTTGAATATGATGAGCTGCTCTCTATGTGGAAATTTGTTAT 481
Qy 634 AAGAGCTCT 642
Db 482 TAAATTTCT 490

RESULT 21
US-08-370-648-18
; Sequence 18, Application US/08370648
; Patent No. 5648226
; GENERAL INFORMATION:
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry;
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,648
; FILING DATE: 10-JANUARY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-JULY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5648226man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5323.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200

FILING DATE: 21-September-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/370,648
FILING DATE: 10-January-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250,162
FILING DATE: 27-May-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/096,039
FILING DATE: 22-July-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5858689man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5323.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 539 base pairs
TYPE: nucleic acid
STRADEDNESS: single
TOPOLOGY: linear

US-08-531-662B-18

Query Match 25.4%; Score 169.8; DB 2; Length 539;
Best Local Similarity 67.1%; Pred.No.1.8e-44;
Matches 288; Conservative 0; Mismatches 132; Indels 9; Gaps 3;

QY 217 GTGAATAATGATGGCGAGGAAGATCAAC--ATATAGGCCTAGGCCAAGAAGAAAGTTT 273
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
75 GTGAATAATGATGGCGAGGAGATCGACCTATTATGTGGCCTAGACCAGGCGCTATGT 134
QY 274 ACAGCTCCTGAGCTGATGGGGCTATGCTTACTGGCTCCCTTTGTCCCAGGAACCCAC 333
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
135 ACAGCTCCTGAAGTGATGGGGCTATGCGGCC---CGAGCAGTTCAGTGATGAAGTGG 191
QY 334 TGATGAAGAGCTTAAGAAGAGAACACCCACCTAAAAGTCGGAACTCTACACCTGATCA 393
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
192 ACCACAACCTGNAGAGGGGAACAGCAACTCAACGTCAGGATCTCTGAGCTGCTCA 251
QY 394 GAAGAGAGAGATGATCAGGTGCGAGCTGAGATCAAGTGCTGNCTGGAGCGCATCT 453
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
252 GGAGGAGAGGATG---AGGGAGCATCTGCAAGTCAAGGGCCGAAGCTGAAGCTGATG 308
QY 454 CCAGGAGCTATCTCAGACAAACACTGGGGATGGATGTGAAGTGGTACTGATGTCAAGGG 513
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
309 CARGAACAGGCTCACCCACAGACTGGGTGTGAGTGTGAGATGCTGATGGGSCAGGA 368
QY 514 GAAGATTTCACAAAGCAGACCTTTAAATGTCAGAGCAGGTGAGGGAATCACA 573
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
369 GGTGGACCGCCCAAATCCAGAGGAGGTGAACAACGCTGAAGAGGTGAAAGCAATCACA 428
QY 574 GGTTTAAGAGAGATAAGCTGAAAAACACAAACTGTTTTATATTAGATATTTTACTTT 633
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
429 GTGTTAAGAGAGACAGTTGAAATGATGTCAGGCTGCTCTATGTTGGAATTTGTTTCA 488
QY 634 AAAGATCT 642
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
489 TAAATTTCT 497

RESULT 23
US-08-669-161A-18
Sequence 18, Application US/08669161A
Patent No. 603481
GENERAL INFORMATION:
APPLICANT: DeBacker, Olivier; Van den Eynde,
APPLICANT: Benoit; Boon-Falleur, Thierry
TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen
TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof

QY 394 GAAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAGTGGCTGCTGACCTGGAAGCCGATCT 453
Db 252 GGAGGGAGAGGATG---AGGGAGCATCTGCAGGTCAAGGGCCGGAAGCCCTGGAAGCTGTAG 308
QY 454 CCAGGAGCTATGTCAGACAAAGACTGGGATGGATGGAAGTGGTACTGTGATGTCAAGGG 513
Db 309 CCAGGAACAGGGTCAACCCAGACTGGGTGTGAGTGTGAAGTGTCTCTGATGGCGGGA 368
QY 514 GAAGATTCACCAAAAGCAGAGACATTTAAATGCCAGAGCAGAGGTGAAGGGAATACCA 573
Db 369 GGTGGACCCGCCCAATCCAGAGAGGTGAAACGCCCTGAAGAAGGTGAAAGCAATACCA 428
QY 574 GGTTTAAAGGAGATAGCTGAAACACACACACTGTTTTTATATTAGATATTTTACTTT 633
Db 429 GTGTTAAAGAGACACCTTGAATGATGACGCTCTCTATGTTGGAAATTTGTTCA 488
QY 634 AAAGAGTCT 642
Db 489 TAAATTTCT 497
RESULT 25
US-08-370-648-16
; Sequence 16, Application US/08370648
; Patent No. 5648226
; GENERAL INFORMATION:
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry;
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,648
; FILING DATE: 10-JANUARY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-JULY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5648226man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5323.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3684
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 540 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-370-648-16
Query Match 25.4%; Score 169.8; DB 1; Length 540;
Best Local Similarity 67.1%; Pred. No. 1.8e-44;
Matches 288; Conservative 0; Mismatches 132; Indels 9; Gaps 3;

QY 217 GTGAATATTCAGTGGCGGAGAGATCAAC---ATATAGCCCTAGGCCAGGAGAGAGTTT 273
Db 76 GTGAATATTCAGTGGCGGAGAGATCGACCTATTATTGSCCTAGACCAAGGCGCTATGT 135
QY 274 ACAGCTCTCTGAGCTGATTTGGGCTATGCTTACTGGCTCCCTTTTCTCCAGAAACCCAC 333
Db 136 ACAGCTCTCTGAAATGATTTGGGCTATGCGGCC---CGAGCAGTTCAAGTGAAGTGA 192
QY 334 TGATGAAGACCTTAAAGAGAGAAACCCACCTATAAGTCGGAATCCCTACACCTGATCA 393
Db 193 ACCAGCAACACCTGAAGAGAGGGAACCAAGCACTCAACGTCAGATCTCTCAGCTGCTCA 252
QY 394 GAAAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAGTGGCTTACCTGACCTGGAAGCCGATCT 453
Db 253 GGAGGGAGAGGATG---AGGGAGCATCTGCAGGTCAAGGGCCGGAAGCTGAAGCTGATAG 309
QY 454 CCAGAGCTATGTCAGACAAAGACTGGGATGGATGGAAGTGGTACTGTAAGGTGGTACTGTCAGGG 513
Db 310 CCAGGAACAGGGTCAACCCAGACTGGGTGTGAGTGTGAAGATGGTCTCTATGTTGGCAAGGA 369
QY 514 GAAGATTCACCAAAAGCAGAGACATTTAAATGCCAGAGCAGAGGTGAAGGGAATACCA 573
Db 370 GATGACCCGCCCAATCCAGAGAGGTGAAACGCCCTGAAGAAGGTGAAAGCAATACCA 429
QY 574 GGTTTAAAGGAGATAGCTGAAACACACACTGTTTTTATATTAGATATTTTACTTT 633
Db 430 GTGTTAAAGAGAGGACACCTTGAATGATGACGCTCTCTATGTTGGAAATTTGTTCA 489
QY 634 AAAGAGTCT 642
Db 490 TAAATTTCT 498
RESULT 26
US-08-531-662B-16
; Sequence 16, Application US/08531662B
; Patent No. 5858689
; GENERAL INFORMATION:
; APPLICANT: Van der Bruggen, Pierre;
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: Isolated, Truncated Nucleic
; TITLE OF INVENTION: Acid Molecules Which Code For GAGE Tumor
; TITLE OF INVENTION: Rejection Antigen, The Tumor Rejection Antigen,
; TITLE OF INVENTION: and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,662B
; FILING DATE: 21-September-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/370,648
; FILING DATE: 10-January-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-May-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-July-1993
; ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. 585868man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5323.3
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 540 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-531-662B-16

Query Match 25.4%; Score 169.8; DB 2; Length 540;
Best Local Similarity 67.1%; Pred. No. 1.8e-44;
Matches 288; Conservative 0; Mismatches 132; Indels 9; Gaps 3;

217 GTGAATATGATGTTGGCGAGAGATCAAC---ATATAGCCTAGGCCCAAGAAAGTTT 273
Db GTGAATATGATGTTGGCGAGAGATCAACCTATTATTGGCTTAGACCAAGCGCTATGT 135

274 ACAGCCTCTGAGTGTGTTGGGCTATGTTACTGGCTCCCTTTCTCCAGGAACCCAC 333
Db ACAGCCTCTGAGTGTGTTGGGCTATGTTACTGGCTCCCTTTCTCCAGGAACCCAC 192

334 TGATGAAGACCTTAAGAGAGAGAAACCCACCTAAAGTTCGAATCTTACCTGATCA 393
Db TGATGAAGACCTTAAGAGAGAGAAACCCACCTAAAGTTCGAATCTTACCTGATCA 252

394 GAAGAGAGAGATGATCAGGTCAGCTGAGATGTAAGTGGTGTGTAAGTGTGTAAGT 453
Db GAAGAGAGAGATGATCAGGTCAGCTGAGATGTAAGTGGTGTGTAAGTGTGTAAGT 309

454 CCAGGAGCTATGTCAGACAAAGACTGGGGATGGATGTGAAGTGGTGTGTAAGTGG 513
Db CCAGGAGCTATGTCAGACAAAGACTGGGGATGGATGTGAAGTGGTGTGTAAGTGG 369

514 GAAGATTTCTACCAAGACAGACACTTTAAATGCCAGAGCAGGTGAAGGGAATCA 573
Db GAAGATTTCTACCAAGACAGACACTTTAAATGCCAGAGCAGGTGAAGGGAATCA 429

574 GGTTTAAAGAGATAGTGAACCAACCAACCAACCAACCAACCAACCAACCAACCA 633
Db GGTTTAAAGAGATAGTGAACCAACCAACCAACCAACCAACCAACCAACCAACCA 489

634 AAAGATCT 642
490 TAAATTTCT 498

RESULT 27
US-08-669-161A-16
Sequence 16, Application US/08669161A
Patent No. 6013481
GENERAL INFORMATION:
APPLICANT: DeBacker, Olivier; Van den Eynde,
APPLICANT: Benoit; Boon-Talieu, Thierry
TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/669,161A
FILING DATE: 24-June-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/531,662
FILING DATE: 21-September-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/370,648
FILING DATE: 10-January-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250,162
FILING DATE: 27-May-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/096,039
FILING DATE: 22-July-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6013481man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5443
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 16:
LENGTH: 540 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-669-161A-16

Query Match 25.4%; Score 169.8; DB 3; Length 540;
Best Local Similarity 67.1%; Pred. No. 1.8e-44;
Matches 288; Conservative 0; Mismatches 132; Indels 9; Gaps 3;

217 GTGAATATGATGTTGGCGAGAGATCAAC---ATATAGCCTAGGCCCAAGAAAGTTT 273
Db GTGAATATGATGTTGGCGAGAGATCAACCTATTATTGGCTTAGACCAAGCGCTATGT 135

274 ACAGCCTCTGAGTGTGTTGGGCTATGTTACTGGCTCCCTTTCTCCAGGAACCCAC 333
Db ACAGCCTCTGAGTGTGTTGGGCTATGTTACTGGCTCCCTTTCTCCAGGAACCCAC 192

334 TGATGAAGACCTTAAGAGAGAGAAACCCACCTAAAGTTCGAATCTTACCTGATCA 393
Db TGATGAAGACCTTAAGAGAGAGAAACCCACCTAAAGTTCGAATCTTACCTGATCA 252

394 GAAGAGAGAGATGATCAGGTCAGCTGAGATGTAAGTGGTGTGTAAGTGTGTAAGT 453
Db GAAGAGAGAGATGATCAGGTCAGCTGAGATGTAAGTGGTGTGTAAGTGTGTAAGT 309

454 CCAGGAGCTATGTCAGACAAAGACTGGGGATGGATGTGAAGTGGTGTGTAAGTGG 513
Db CCAGGAGCTATGTCAGACAAAGACTGGGGATGGATGTGAAGTGGTGTGTAAGTGG 369

514 GAAGATTTCTACCAAGACAGACACTTTAAATGCCAGAGCAGGTGAAGGGAATCA 573
Db GAAGATTTCTACCAAGACAGACACTTTAAATGCCAGAGCAGGTGAAGGGAATCA 429

574 GGTTTAAAGAGATAGTGAACCAACCAACCAACCAACCAACCAACCAACCAACCA 633
Db GGTTTAAAGAGATAGTGAACCAACCAACCAACCAACCAACCAACCAACCAACCA 489

634 AAAGATCT 642
490 TAAATTTCT 498

RESULT 28
US-08-602-039-16
Sequence 16, Application US/08602039
Patent No. 6069001

GENERAL INFORMATION:
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry;
; TITLE OF INVENTION: Isolated, truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,039
; FILING DATE: 15-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-JULY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No 6069001man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5323.2
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 540 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-602-039-16
Query Match 25.4%; Score 169.8; DB 3; Length 540;
Best Local Similarity 67.1%; Pred. No. 1.8e-44;
Matches 288; Conservative 0; Mismatches 132; Indels 9; Gaps 3;
QY 217 GTGAAATATGAGTTGGCGAGGAGATCAAC---ATATAGGCTTAGCCCAAGAAAGTTT 273
DB 76 GTGAAATATGAGTTGGCGAGGAGATCGACCTATTATTGGCTTAGACCAAGCGCTATGT 135
QY 274 ACAGCTCTGAGCTGATGGGCTATGTTACTGGCTCCCTTTGTCCCAAGAACCCAC 333
DB 136 ACAGCTCTGAAATGATTGGGCTATGGGCC---CGAGCAGTTCAAGTGAAGTGA 192
QY 334 TGATGAAGACCTTAAGAGAGAGAAACCCACCTAAAGTCGGAATCCCTACACCTGATCA 393
DB 193 ACCAGCAACACCTGAAGAGAGGGAACACGACCACTCAAGCTCAGGATCCTGCAGCTGCTCA 252
QY 394 GAAGAGAGAGATGATCAGGGTCAGCTGAGATTCAGTGCCTGACCTGGAGCCGATCT 453
DB 253 GGAGGGAGAGATG---AGGGAGATCTGAGGTCAGGGCCGAGCCCTGAGCTGATAG 309
QY 454 CCAGGAGCTATGTCAGACAAAGACTGGGGATGGATGTGAAGTGGTGTACTGATCAAGGG 513
DB 310 CCAGGACAGGGTCACCCACAGACTGGTGTGAGTGTGAAGATGGTCTGATGGCAGGA 369
QY 514 GAAGATTCACCAAAACAGAGCACTTTAAATGCCAGAGCAGGTCGAGGGAATCACA 573
DB 370 GATGGACCCGCAATATCCAGAGGAGGTGAAACCGCTGAAAGAGGTTGAAAGCAATCACA 429

QY 574 GGTTTAAAGGAGATAGCTCAAAACACACAAACCTGTTTTATATTAGATATTTTACTTT 633
DB 430 GTGTTAAAGAGGACCGCTTGAATGATGCAGGCTGCTCTATGTTGAAATTTTTCAT 489
QY 634 AAGAGTCT 642
DB 490 TAAATTTCT 498
RESULT 29
US-09-163-748C-2
; Sequence 2, Application US/09163748C
; Patent No. 6509172
; GENERAL INFORMATION:
; APPLICANT: DeBacker, Olivier
; APPLICANT: Van den Eynde, Benoit
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Which Encode Member Of The Gage
; TITLE OF INVENTION: Family, The Proteins Encoded, And Uses Thereof
; FILE REFERENCE: LUD 558
; CURRENT APPLICATION NUMBER: US/09/163,748C
; CURRENT FILING DATE: 1998-09-30
; NUMBER OF SEQ ID NOS: 27
; SEQ ID NO 2
; LENGTH: 526
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; US-09-163-748C-2
Query Match 25.1%; Score 168.2; DB 4; Length 526;
Best Local Similarity 66.9%; Pred. No. 5.7e-44;
Matches 287; Conservative 0; Mismatches 133; Indels 9; Gaps 3;
QY 217 GTCAATATGAGTTGGCGAGGAGATCAAC---ATATAGGCTTAGCCCAAGAAAGTTT 273
DB 48 GTCAATATGAGTTGGCGAGGAGATCGACCTATTATTGGCTTAGACCAAGCGCTATCT 107
QY 274 ACAGCTCTGAGCTGATGGGCTATGTTACTGGCTCCCTTTGTCCCAAGAACCCAC 333
DB 108 ACAGCTCTGAAATGATTGGGCTATGGGCC---CGAGCAGTTTCAGTGAAGTGA 164
QY 334 TGATGAAGACCTTAAGAGAGAGAAACCCACCTAAAGTCGGAATCCCTACACCTGATCA 393
DB 165 ACCAGCAACACCTGAAGAGAGGGAACCGCACTCAAGCTCAGGATCCTGCAGCTGCTCA 224
QY 394 GAAGAGAGAGATGATCAGGGTCAGCTGAGATTCAGTGCCTGACCTGGAAGCCGATCT 453
DB 225 GGAGGGAGAGATG---AGGGAGATCTGCAGGTCAGGGCCGAGCCCTGAAGCTCATAG 281
QY 454 CCAGGAGCTATGTCAGACAAAGACTGGGATGGATGTGAAGTGGTACTGATCTCAAGGG 513
DB 282 CCAGGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTGATGGCAGGA 341
QY 514 GAAGATTCACCAAAACAGAGCACTTTAAATGCCAGAGAGGTTGAAAGAGGTTGAAAGTCA 573
DB 342 GATGACCCGCCAATCCAGAGAGGTTGAAACCGCTGAAAGAGGTTGAAAGCAATCACA 401
QY 574 GGTTTAAAGGAGATAGCTGAAACACACAAACTGTTTTATATTAGATATTTTACTTT 633
DB 402 GTGTTAAAGAGGACCGCTTGAATGATGCAGGCTGCTCTATGTTGAAATTTTTCAT 461
QY 634 AAGAGTCT 642
DB 462 TAAATTTCT 470
RESULT 30
US-08-370-648-15
; Sequence 15, Application US/08370648
; Patent No. 5648226
; GENERAL INFORMATION:

APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
 APPLICANT: Boon-Falleur, Thierry;
 TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
 TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
 TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
 NUMBER OF SEQUENCES: 18
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Felfe & Lynch
 STREET: 805 Third Avenue
 CITY: New York City
 STATE: New York
 COUNTRY: USA
 ZIP: 10022
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: PC-DOS
 SOFTWARE: Wordperfect
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/370,648
 FILING DATE: 10-JANUARY-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/250,162
 FILING DATE: 27-MAY-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/096,039
 FILING DATE: 22-JULY-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Hanson, No. 5648226man D.
 REGISTRATION NUMBER: 30,946
 REFERENCE/DOCKET NUMBER: LUD 5323.2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 688-9200
 TELEFAX: (212) 838-3884
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 560 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-370-648-15

Query Match 24.2%; Score 162.2; DB 1; Length 560;

Best Local Similarity Matches	62.4%; Pred. No. 4.9e-42; 289; Conservative 0; Mismatches 168; Indels 6; Gaps 2;
Qy	180 ATAGGTCGCAGGCCAGTCTCCCGACGAACTGAATAATAGTAGTGCGCAGGAA 239
Db	62 ATAGGTGGATCTTCTCTGCCAACTCATATTTCACACAGATGAATCTCAGTAGAGAANAAT 121
Qy	240 GATCAACATATAGSCCTTAGGCCAACAGATTTTACAGCCTCTTGAGCTGATTGGGGCTA 299
Db	122 CGACTTATTATGGCTCTAGACAAGCGCCTATGTACAGCCTCTCTGAAGTGAATTGGGCTTA 181
Qy	300 TGTCTACTTGGCTCCOCTTTGTCCCGAGCAACCCTGATGAAGAGCCTAAAGAAGAGAAAC 359
Db	182 TCGCGCCCGAG-- --CAGTTCAGTGATGAAGTGGAAACCAGCAACACCTTGAAGAAGGGGAAC 238
Qy	360 CACCCACTAAAAGTCGGAAATCCTCACACCTGATCAGAGAGAGAGATGATCAGGGTGCAG 419
Db	239 CAGCAACTCAACGTCAGGATCCTGCAGCTGTCTCAGGAGGGAGAGGATG---AGGGAGCAT 295
Qy	420 CTGAGATTCAGTGCCTGACCTGGAGGCCGATCTCCAGGAGCTATGTCAGACAAGACCTG 479
Db	296 CTGCAGGTC AAGGGCCGAAGCCTGAGTGTATAGCCAGGAACAGGGTCACCCACAGACTG 355
Qy	480 GGGATGGATGTGAAGGTGGTACTGATGTCAAGGGGGAAGATTCTACCAAAGCCAGAGCACT 539
Db	356 GGTGTGAGTGTGAAGATGGTCTCTGATGGCAGGAGATGGAACCGCCCAATCCAGAGGAGG 415
Qy	540 TTAAAAATCCAGAACGAGGTGAAGGGAAATCACAGGTTTTAAGGAAGATATAGCTGAAACA 599

Db 416 TGAAGAACGCTGACAGAGGTGAAAAAGCAANTACACAGTCTTTAAAGAGGCACGTTGAAATG 47

Qy 600 ACACAAACTGTTTTATATTAGATATTTTACTTTTAAAGAGTCT 642

Db 476 ATGCAGGCTGCTCCTATGTTGAAATTTGTTTCATTAAATCT 518

RESULT 31

US-08-531-662B-15

Sequence 15, Application US/08531662B

Patent No. 5856869

GENERAL INFORMATION:

APPLICANT: Van der Bruggen, Pierre;

APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;

APPLICANT: Boon-Falleur, Thierry

TITLE OF INVENTION: Isolated, Truncated Nucleic

TITLE OF INVENTION: Acid Molecules Which Code For GAGE Tumor

TITLE OF INVENTION: Rejection Antigen, The Tumor Rejection Antigen,

TITLE OF INVENTION: and Uses Thereof

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felfe & Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/531,662B

FILING DATE: 21-September-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/370,648

FILING DATE: 10-January-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/250,162

FILING DATE: 27-May-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/096,039

FILING DATE: 22-July-1993

ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. 5856869man D.

REGISTRATION NUMBER: 30,946

REFERENCE/DOCKET NUMBER: LUD 5323.3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 560 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-531-662B-15

Query Match
24.2%; Score 162.2; DB 2; Length 560;

	Best Local Similarity	62.4%;	Pred. No. 4.9e-42;	Mismatches	169;	Indels	6;	Gaps	2
	Matches	289;	Conservative	0;					
QY	180	ATAGGTCGCGAGGGCAGTCTCC	CAGGNACTGAATACTGAATAAT	TAGTGAATATGAGTTGGCGAGAA	239				
Db	62	ATAGTTCGATCTTCTCGCCAACTCATATTT	CACACAGATGAATCTCAGTAGAGAAAAT	121					
QY	240	GATCAACATATAGGCTTAGGCCAAGAAAGTTT	CACGCTCCTGAGCTGATTTGGGGGCTA	299					
Db	122	CGACTATTATTGGCTTAGACRAGGCGCTAT	GTACAGCTCCTGAGTGTGTTGGGCTA	181					
QY	300	TGCTTACTGGTCCGCTTTTGTC	CCAGNAACCCACTGTAGAGAGCGCTTAAGAGAGAGAAC	359					

```

182 TCGGCCCCGAG---CAGTTCAGTGATGAAGTGGACACGACACCTTGAGAGGGGAAC 238
360 CACCACCTAAAGTCGGAAATCCTACCTGATCAGAAGAGAGAGATGATCAGGGTCCAG 419
239 CAGCAACTCAACGTCAGGATCTCAGCTGCTCAGGAGGGAGAGATG---AGGGAGCAT 295
420 CTGAGATTCAAGTGGCTCCTCAGCTGGAAGCCGATCTCCAGGAGCTATGTCCAGACAAAGACTG 479
296 CTGCAAGTCAGAGCCGAGGCTGAGCTGATGATGACGAGGACAGGGTCCACCCAGACTG 355
480 GGGATGATGTAAGTGGTACTGATGTCAGAGGGAAGATTCCTACCAAAAGCAGAGCACT 539
356 GGTGTGAGTGAAGATGGTCTGATGGCGAGGAGATGGACCCGCCAAATCCAGAGGAGG 415
540 TTAATAATCCAGAGCAGCTGAGGGAATCAGAGGTTTAAAGGAGATGATGATGAAACA 599
416 TGAATAACGCTGAGAGAGGTGAAGCAATCAGTGTGTTAAAGAGCAGCTTGAATG 475
600 ACACAACTGTTTATATAGATATTTTACTTTAAAGAGTCT 642
476 ATGCAGGCTGCTCTATGTTGAAATTTGTTCAATTAATTTCT 518

```

RESULT 32

```

US-08-669-161A-15
; Sequence 15, Application US/08669161A
; Patent No. 6013481
; GENERAL INFORMATION:
; APPLICANT: DeBacker, Olivier; Van den Eynde,
; APPLICANT: Benoit; Boon-Falleur, Thierry
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/669,161A
; FILING DATE: 24-June-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/531,662
; FILING DATE: 21-September-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/370,648
; FILING DATE: 10-January-1995
; APPLICATION DATA:
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-May-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-July-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6013481man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5443
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 560 base pairs

```

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-669-161A-15
Query Match 24.2%; Score 162.2; DB 3; Length 560;
Best Local Similarity 62.4%; Pred. No. 4,9e-42;
Matches 289; Conservative 0; Mismatches 168; Indels 6; Gaps 2;
QY 180 ATAGGTCCCGAGGCGCAGTCTCCAGGAGCACTGAAATATGTAATATGATGTTGGCCAGGAA 239
Db 62 ATAGGTGATCTCTCTCGCCAACTCATATTTTCCACACAGATGAATCTCAGTAGAGGAAAT 121
QY 240 GATCAACATATAGGCTTAGGCCCAAGAAAGTTTACAGCTCTCCAGCTGATTTGGGGCTA 299
Db 122 CGAOCCTATTATTGGCTTAGACCAAGCGCTATGTACAGCTCTCTGAGTGTGGSCCTA 181
QY 300 TGTCTACTGGCTCCCTCTTGTCCAGGAAACCCACCTGATGAAAGAGCTTAAAGAGAGAAAC 359
Db 182 TCGGCCCCGAG---CAGTTTCAGTGATGAAGTGGAAACCCAGCAACCTGAAAGAGGGAAC 238
QY 360 CACCACCTAAAGTCGGAAATCCTACCTGATCAGAAGAGAGAGATGATCAGGGTGCAG 419
Db 239 CAGCAACTCAACGTCAGGATCTCTGAGTGTGTCAGAGGAGAGATG---AGGAGAT 295
QY 420 CTGAGATTCAAGTGGCTGACCTGGAAGCCGATCTCCAGGAGCTATGTCCAGACAAAGACTG 479
Db 296 CTGCAAGTCAGAGCCGAGGCTGAGCTGATGACGAGGACAGGGTCCACCCAGACTG 355
QY 480 GGGATGATGTAAGTGGTACTGATGTCAAGGGGAAGATTCCTACCAAAAGCAGAGCACT 539
Db 356 GGTGTGAGTGAAGATGGTCTGATGGCGAGGAGATGGACCCGCCAAATCCAGAGGAGG 415
QY 540 TTAATAATCCAGAGCAGCTGAGGGAATCAGAGGTTTAAAGGAGATGATGATGAAACA 599
Db 416 TGAATAACGCTGAGAGAGGTGAAGCAATCAGTGTGTTAAAGAGCAGCTTGAATG 475
QY 600 ACACAACTGTTTATATAGATATTTTACTTTAAAGAGTCT 642
Db 476 ATGCAGGCTGCTCTATGTTGAAATTTGTTCAATTAATTTCT 518

```

RESULT 33

```

US-08-602-039-15
; Sequence 15, Application US/08602039
; Patent No. 6069001
; GENERAL INFORMATION:
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry;
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,039
; FILING DATE: 15-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/370,648
; FILING DATE: 10-JANUARY-1995
; APPLICATION NUMBER: 08/250,162

```


FILING DATE: 27-MAY-1994
PRIOR APPLICATION NUMBER: 08/096,039
APPLICATION DATE: 22-JULY-1993
FILING DATE: 22-JULY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6069001man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5323.2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 560 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-602-039-15

Query Match 24.2%; Score 162.2; DB 3; Length 560;
Best Local Similarity 62.4%; Pred. No. 4.9e-42;
Matches 289; Conservative 0; Mismatches 168; Indels 6; Gaps 2;

QY 180 ATAGGTCGGCAGGCGAGTCTCCAGGAAGTGAATAGTGAATATGAGTTGCGGAGAA 239
DB 62 ATAGGTCGATCTTCTCGCCAACTCATATTTACACAGATGATCTCAGTAGGGAAT 121
QY 240 GATCAACATATAGCGCTAGCGCCAGAGAAAGTTTACAGCCCTCTGAGCTGATTTGGGCTA 299
DB 122 CGACCTATTATTGGCGCTAGACCAAGGCGCTATGTACAGCCCTCTGAGTGAATTTGGGCTA 181
QY 300 TGCTTACTGCTCCCTTTCTCCAGGAAGCCACTGATGAGAGCCCTAAAGAGAGAAAC 359
DB 182 TGGGCCCCGAG--CAGTTGATGATGAGTGAACACAGCAACCTGAGAGAGGGGAC 238
QY 360 CACCACATAAAGTCGGAATCTTACACCTGATCAGAGAGAGAGATGATCAGGGTGCAG 419
DB 239 CAGCAACTCAAGCTCAGGATCTCTGAGTCTGCTCAGGAGGAGAGGATG---AGGAGCAT 295
QY 420 CTGAGATTCAAGTCCCTGACCTGGAAGCCGATCTCCAGGAGCTATCTCAGACAAAGCATG 479
DB 296 CTGAGTCAAGGCGGAGGCTGAGCTGATGACGAGGAGGCTACCCACAGACTG 355
QY 480 GGGATGATGTAGAGTGGTACTGATGTCAGAGGGGAGATTTTACCAAAAGCAGAGCACT 539
DB 356 GGTGTGAGTGTGAAGATGCTCTGATGGCAGGAGATGGACCCGCCCAATCCAGAGGAGG 415
QY 540 TTAATGTCAGAGCAGGTCAGGGAAGTCAAGGCTTAAAGAGAGATGATGAGTGAACA 599
DB 416 TGAAGAGCCCTGAGAGGTTGAAAGCAATCAGAGTGTAAAGAGAGGCGACGTTGAAATG 475
QY 600 ACACAACTGTTTATTATAGATATTTTACTTTAAAGAGTCT 642
DB 476 ATGAGGCTGCTCTATGTTGGAAATTTGTTCAATTAATTTCT 518

RESULT 34

US-09-702-705-1827
Sequence 1827, Application US/09702705
Patent No. 6504010

GENERAL INFORMATION:
APPLICANT: Wang, Tongtong

APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.

APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom

APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc

APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.478C14

CURRENT APPLICATION NUMBER: US/09/702,705
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 1827
LENGTH: 288
TYPE: DNA
ORGANISM: Homo sapiens
US-09-702-705-1827

Query Match 23.2%; Score 155.2; DB 4; Length 288;

Best Local Similarity 77.3%; Pred. No. 5.9e-40;
Matches 201; Conservative 0; Mismatches 58; Indels 1; Gaps 1;

QY 319 GTCCAGGAACCCACTGATGAGAGAGCTTAAAGAGAGAAACCCCACTAAAGTCGAA 378
DB 29 GTCCAGGAGCCAGTATGAGAGAGCCCAAAAGAGAGACAGCAGCTGAAGTCGGGA 88

QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGA 438
DB 89 TCCTACACCTGGCAGCAGACAGAGAGATCAGGATACAGCTGAGATCCCAAGTCGCGGA 148

QY 439 CCTGGAAGCCGATCTCCAGGAGCTATGTACAGACAAAGACTGGGGATGGATGTGAAGT-G 497
DB 149 CATGGAAGTGTATCTGCAAGAGCTGCATCACTCAACACCGGGGATTAATCTGATTGG 208

QY 498 GTACTGATGTCAAGGGGAGATTTACCAAAAGCAGAGCACTTTAAATGCCAGAGCAG 557
DB 209 GTTCCGGGCTCAAGGTGAAGATAATACCTAAAGAGGAACACTGTAAATGCCAGAGCAG 268

QY 558 GTGAAGGGAAATCAGGTT 577
DB 269 GTGAAGAGCAACCAAGTT 288

RESULT 35

US-09-736-457-1827
Sequence 1827, Application US/09736457
Patent No. 6509448

GENERAL INFORMATION:
APPLICANT: Wang, Tongtong

APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.

APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom

APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc

APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun

APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13

NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 1827

LENGTH: 288
TYPE: DNA
ORGANISM: Homo sapiens
US-09-736-457-1827

Query Match 23.2%; Score 155.2; DB 4; Length 288;

Best Local Similarity 77.3%; Pred. No. 5.9e-40;
Matches 201; Conservative 0; Mismatches 58; Indels 1; Gaps 1;

QY 319 GTCCAGGAACCCACTGATGAGAGAGCTTAAAGAGAGAAACCCCACTAAAGTCGAA 378
DB 29 GTCCAGGAGCCAGTATGAGAGAGCCCAAAAGAGAGACAGCAGCTGAAGTCGGGA 88

QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGA 438
DB 89 TCCTACACCTGGCAGCAGACAGAGAGATCAGGATACAGCTGAGATCCCAAGTCGCGGA 148

Db	89	TCCTACACCTGGCAGCAGACAGAGATCAGGATACAGCTGAGATCCAGTGGCGA	148
Qy	439	CTGGAAGCCGATCTCCAGAGCTATGTGACAGAAACACTGGGGATGGATGTGAAGGT-G	497
Db	149	CATGGAAGGTGATCTGCAAGAGCTGATCAGTCAAAACACGGGGATAAATCTGGATTTGG	208
Qy	498	GTACTGATGTCAGGGGGAAGATTCTACCAAAGCAGAGCACTTTAAATGCCAGAGACAG	557
Db	209	GTTCGGCGTCAAGGTGAAGATTAATACCTAAGAGGACACTGTAAATGCCAGAGACAG	268
Qy	558	GTGAAGGGAATCACAGTTT	577
Db	269	GTGAAGAGCAACACAGTTT	288
RESULT 36			
US-09-328-475C-263			
; Sequence 263, Application US/09328475C			
; Patent No. 6476207			
; GENERAL INFORMATION:			
; APPLICANT: Zhang, Jimmy			
; APPLICANT: Astel, Jon H.			
; APPLICANT: Carroll III, Eddie			
; APPLICANT: Endege, Wilson O.			
; APPLICANT: Ford, Donna M.			
; APPLICANT: Monahan, John E.			
; APPLICANT: Schlegel, Robert			
; APPLICANT: Steinmann, Kathleen E.			
; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT			
; TITLE OF INVENTION: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER			
; FILE REFERENCE: 1532.002/200130.463			
; CURRENT APPLICATION NUMBER: US/09/328.475C			
; CURRENT FILING DATE: 1999-06-09			
; NUMBER OF SEQ ID NOS: 341			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 263			
; LENGTH: 656			
; TYPE: DNA			
; ORGANISM: Homo Sapien			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (1)...(656)			
; OTHER INFORMATION: n = A,T,C or G			
US-09-328-475C-263			
Query Match 22.2%; Score 148.6; DB 4; Length 656;			
Best Local Similarity 62.0%; Pred. No. 1.2e-37;			
Matches 291; Conservative 0; Mismatches 136; Indels 42; Gaps 2;			
Qy	90	TGCCACCTGGTGACCGCTGGGGAGCTGTGAGAGTGTGAGGGGACGTTCCAGCCGCTG	149
Db	82	TACCGCGGGGAGAACGCCAGGAGCTGTGAGAGTGTGAGTGGGTTCTCTGTCTCGG	141
Qy	150	GACTCTTTCTCTCTCTAGAGCGGCGCTATAGTCCGAGGGCAGTCTCCAGGAC	209
Db	142	ACACTTTTTTCTCTACTGAGACTCATCTGGTAGATCCGAGGGCAGTCTCCAGGGGC	201
Qy	210	TCAAAATAGTGAATATGATGTGGCAGGAAGATCAACATATAGGCTAGGCCAAGAAGAA	269
Db	202	TGAAGTTGGAATATGGTTTCTAAGAGATTAACTATCGGGGTAGACCAATGATCT	261
Qy	270	GTTTAAGCCTCCTGAGCTGATTTGGGCTATGCTTACTGTGCTCCCTTTGTCCAGGAC	329
Db	262	ATGTAGAATCTTCTG-----AGGAGT	282
Qy	330	CCACTGATCAAGAGCCTAAGAGAGAAACCCACCTAAAGTCGGATCTACACCTG	389
Db	283	CCAGTATGAGCAACTGACGAGTGGATCAACCACTCAAGTCAGATTCTACACCTG	342
Qy	390	ATCAGAAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTCTGACCTGGAAGCCG	449
Db	343	CTGAAGAGAGAGAGGATG---AGGAGCATCTGCAGCTCAAGGGCAGGAGCCTGGAAGCTG	399
Qy 450 ATCTCCAGGAGCTATGTTCAGACCAAGACTGGGATGGATGTGAAGTGGTGTCTGATCTCA 509			
Db 400 ATAGCCAGGAGCTGTTTCAGCCAAAGACTGGGTGTGAGCTTGGAGATGGTCTGTATACCA 459			
Qy 510 AGGGGAAGATTCACCAAAAGCAGGACACTTTAAATGCCAGAGACAGG 558			
Db 460 AAGGGGTTGCTCGGAAATGAAGAGCAGATGAACCTGCCCGNAGAGG 508			
RESULT 37			
US-08-370-648-1			
; Sequence 1, Application US/08370648			
; Patent No. 5848226			
; GENERAL INFORMATION:			
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;			
; APPLICANT: Boon-Falleur, Thierry;			
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid			
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,			
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof			
; NUMBER OF SEQUENCES: 18			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Felfe & Lynch			
; STREET: 805 Third Avenue			
; CITY: New York City			
; STATE: New York			
; COUNTRY: USA			
; ZIP: 10022			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage			
; COMPUTER: IBM PS/2			
; OPERATING SYSTEM: PC-DOS			
; SOFTWARE: Wordperfect			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/370,648			
; FILING DATE: 10-JANUARY-1995			
; CLASSIFICATION: 530			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: 08/250,162			
; FILING DATE: 27-MAY-1994			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: 08/096,039			
; FILING DATE: 22-JULY-1993			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Hanson, No. 5648226man D.			
; REGISTRATION NUMBER: 30,946			
; REFERENCE/DOCKET NUMBER: LUD 5323.2			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (212) 688-9200			
; TELEFAX: (212) 838-3884			
; INFORMATION FOR SEQ ID NO: 1:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 646 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
US-08-370-648-1			
Query Match 21.6%; Score 144.2; DB 1; Length 646;			
Best Local Similarity 68.0%; Pred. No. 3.1e-36;			
Matches 232; Conservative 0; Mismatches 103; Indels 6; Gaps 2;			
Qy	217	GTCAATATGAGTTGGCGAGGAAGATCAACATATAGGCCTAGGCCAAGAAGATTTACA	276
Db	42	GTGAATATGAGTTGGCGAGGAGATCGACCTATCGGCTAGACCAAGACGTACGTAGA	101
Qy	277	GCCTCTGAGCTGAATGGGGCTATGTTACTGGCTCCCTTTGTCCAGGAACCCACTGA	336
Db	102	GCCTCTGAAATGATTGGGCTATGGGC---CGAGCAGTTTCAGTGAATGAAGTGAAC	158
Qy	337	TGAGAGCCTTAAGAGAGAGAACCCCACTAAAGTCGGAATCCCTACACCTGATCAGAA	396
Db	159	AGCAACACCTTGAAGAGAGGAGACCAAGAACTCAACGTAGGATCTCTGAGCTGCTCAGGA	218

```
QY 397 GAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAGCGAGTCTCCA 456
Db 219 GGGAGAGGATG---AGGGAGCATCTGCAGTCAAGGGCCGAAGCCTGAAAGCTGATAGCCA 275
QY 457 GGAGCTATGTCAGACAAAGACTGGGATGGATGTGAAGTGTGACTGATGTCAAGGGAA 516
Db 276 GGAACAGGGTCAACCCAGACTGGGTGTGAGTGTGAAGATGGTCTCTGATGGCAGAGAT 335
QY 517 GATTCTACCAAAGCAGAGCACTTTAAATGCCAGAGCAG 557
Db 336 GGACCCGCCAAATCCAGAGAGGTGAAACGCCCTGAAGAAG 376

RESULT 38
US-08-531-662B-1
; Sequence 1, Application US/08531662B
; Patent No. 5858689
; GENERAL INFORMATION:
; APPLICANT: Van den Bruggen, Pierre;
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: Isolated, Truncated Nucleic
; TITLE OF INVENTION: Acid Molecules Which Code For GAGE Tumor
; TITLE OF INVENTION: Rejection Antigen, The Tumor Rejection Antigen,
; TITLE OF INVENTION: and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531.662B
; FILING DATE: 21-September-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/370,648
; FILING DATE: 10-January-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-May-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-July-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5858689man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5323.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 646 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-531-662B-1
Query Match 21.6%; Score 144.2; DB 2; Length 646;
Best Local Similarity 68.0%; Pred. No. 3.1e-36;
Matches 232; Conservative 0; Mismatches 103; Indels 6; Gaps 2;

QY 217 GTGAATATGATTCGGCAGGAGATCAACATATAGGCTAGGCCAGCAAGAAATTACA 276
Db 42 GTGAATATGATTCGGCAGGAGATCGACCTATCGGCTAGACCAAGACCTACGTAGA 101
```

```
QY 277 GCCTCCTGAGCTGATTCGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAACCCACTGA 336
Db 102 GCCTCCTGAAATGATTCGGGCTATGCGGCC---CGAGCAGTTTCAGTGATGAAGTGGAAACC 158
QY 337 TGAAGAGCCTAAAGAGAGAAACCAACCCACTTAAAGTCGGAATCCTACACTGATCAGAA 396
Db 159 AGCAACACCTGAGAGAGGGGAAACCAAGCAATCAACGTTCAGGATCCTGAGCTGCTCAGGA 218
QY 397 GAGAGAAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAAGCCGATCTCCA 456
Db 219 GGGAGAGGATG---AGGGAGCATCTGCAGGTCAAGGGCCGAAGCCTGAAGCTGATAGCCA 275
QY 457 GGAGCTATGTCAGACAAAGACTGGGATGGATGTGAAGTGTGACTGATGTCAAGGGAA 516
Db 276 GGAACAGGGTCAACCCAGACTGGGTGTGAGTGTGAAGATGGTCTCTGATGGCAGAGAT 335
QY 517 GATTCTACCAAAGCAGAGCACTTTAAATGCCAGAGCAG 557
Db 336 GGACCCGCCAAATCCAGAGAGGTGAAACGCCCTGAAGAAG 376

RESULT 39
US-08-669-161A-1
; Sequence 1, Application US/08669161A
; Patent No. 6013481
; GENERAL INFORMATION:
; APPLICANT: DeBacker, Olivier; Van den Eynde,
; APPLICANT: Benoit; Boon-Falleur, Thierry
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/669.161A
; FILING DATE: 24-June-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/531.662
; FILING DATE: 21-September-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/370,648
; FILING DATE: 10-January-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-May-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-July-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6013481man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5443
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 646 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

US-08-602-039-1
; TOPOLOGY: linear
; US-08-602-039-1
; Query Match 21.6%; Score 144.2; DB 3; Length 646;
; Best Local Similarity 68.0%; Pred. No. 3.1e-36;
; Matches 232; Conservative 0; Mismatches 103; Indels 6; Gaps 2;
; QY 217 GTGAATATGAGTTGGCGAGGAAGATCAACATATAGGCTAGGCGCAAGAGAGAGTTTACA 276
; DB 42 GTGAATATGAGTTGGCGAGGAAGATCAACATATAGGCTAGGCGCAAGAGAGAGTTTACA 101
; QY 277 GCCTCTCAGCTGATGGCGCTATGCTTACTGGCTCCCTTTGTCCTCCAGGAGCCACTGA 336
; DB 102 GCCTCTCAGCTGATGGCGCTATGCTTACTGGCTCCCTTTGTCCTCCAGGAGCCACTGA 158
; QY 337 TGAAGAGCCTTAAGAGAGGAAGAACCAACCACTAAAGTGGGAATCCTACACCTGATCAGAA 396
; DB 159 AGCAACACTGAGAGAGGGAAGAACCAACCACTAAAGTGGGAATCCTACACCTGATCAGAA 218
; QY 397 GAGAGAGATGATCAGAGTGGCGATGATGATGATGATGATGATGATGATGATGATGATGAT 456
; DB 219 GCGAGAGGATG---AGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 275
; QY 457 GAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 516
; DB 276 GGAACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 335
; QY 517 GATTTACCAAG 557
; DB 336 GGACCCGCAATCCAG 376
; RESULT 40
; US-08-602-039-1
; Sequence 1, Application US/08602039
; Patent No. 6069001
; GENERAL INFORMATION:
; APPLICANT: Van den Eynde, Benoît; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry;
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,039
; FILING DATE: 15-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/370,648
; FILING DATE: 10-JANUARY-1995
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-JULY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6069001man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5323.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884

US-08-602-039-1
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 646 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-602-039-1
; Query Match 21.6%; Score 144.2; DB 3; Length 646;
; Best Local Similarity 68.0%; Pred. No. 3.1e-36;
; Matches 232; Conservative 0; Mismatches 103; Indels 6; Gaps 2;
; QY 217 GTGAATATGAGTTGGCGAGGAAGATCAACATATAGGCTAGGCGCAAGAGAGAGTTTACA 276
; DB 42 GTGAATATGAGTTGGCGAGGAAGATCAACATATAGGCTAGGCGCAAGAGAGAGTTTACA 101
; QY 277 GCCTCTCAGCTGATGGCGCTATGCTTACTGGCTCCCTTTGTCCTCCAGGAGCCACTGA 336
; DB 102 GCCTCTCAGCTGATGGCGCTATGCTTACTGGCTCCCTTTGTCCTCCAGGAGCCACTGA 158
; QY 337 TGAAGAGCCTTAAGAGAGGAAGAACCAACCACTAAAGTGGGAATCCTACACCTGATCAGAA 396
; DB 159 AGCAACACTGAGAGAGGGAAGAACCAACCACTAAAGTGGGAATCCTACACCTGATCAGAA 218
; QY 397 GAGAGAGATGATCAGAGTGGCGATGATGATGATGATGATGATGATGATGATGATGATGAT 456
; DB 219 GCGAGAGGATG---AGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 275
; QY 457 GAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 516
; DB 276 GGAACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 335
; QY 517 GATTTACCAAG 557
; DB 336 GGACCCGCAATCCAG 376
; RESULT 41
; US-08-250-162A-1
; Sequence 1, Application US/08250162A
; Patent No. 5610013
; GENERAL INFORMATION:
; APPLICANT: Van den Eynde, Benoît; Boon-Falleur, Thierry
; TITLE OF INVENTION: METHOD FOR DIAGNOSING A DISORDER BY
; TITLE OF INVENTION: DETERMINING EXPRESSION OF GAGE TUMOR REJECTION ANTIGEN PRECUR:
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2 PC-DOS
; OPERATING SYSTEM: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/250,162A
; FILING DATE: 27-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-JULY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5610013man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5323.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884

```
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 648 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-250-162A-1

Query Match      21.6%; Score 144.2; DB 1; Length 648;
Best Local Similarity 68.0%; Pred. No. 3.1e-36;
Matches 232; Conservative 0; Mismatches 103; Indels 6; Gaps 2;

QY 217 GTCAATATCAGTTGGCGAGAGATCAACATATAGCCCTAGCCCAAGAGAGTTTACA 276
DB 44 GTGAATATAGTTGGCGAGAGATCGACCTATCGCCCTAGCCCAAGAGCTGCTAGTA 103
QY 277 GCCTCTGAGCTGATTGGGCTATGCTTACTTGCTCCCTTTGTCCAGGAACCCACTGA 336
DB 104 GCCTCTGAAATGATTGGGCTATGCGGCC---CGAGCAGTTTCAGTGATGAAGTGAACC 160
QY 337 TGAAGAGCTTAAGAGAGAACCCCACTAAAGTCGGAATCCTACACCTGATCAGAA 396
DB 161 AGCAACACCTTGAAGAGGGGAACAGCACTCAAGCTCAGGATCTTGCAGCTGCTCAGGA 220
QY 397 GAGAGAGATGATCAGGCTGAGCTGAGATTCAGTCAAGTGCCTGAGCTGGAAGCCGATCCA 456
DB 221 GCGAGAGGATG---AGGAGCATCTGCAGGTCAAGGCCGAGCCTGAAGCTGATAGCCA 277
QY 457 GAGCTATGTCAGACAAAGATCGGATGATGATGAGTGGTACTGATGTCNAGGGAA 516
DB 278 GGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGTGGTCTGATGGCAGGAGAT 337
QY 517 GATTCTACCAAAAGCAGACACTTTAAATGCCAGAGCAG 557
DB 338 GGACCCGCAATCCAGAGGAGGTGAAGCCCTGAAGAG 378

RESULT 42
US-09-328-475C-78
; Sequence 78, Application US/09328475C
; Patent No. 6476207
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jimmy
; APPLICANT: Astel, Jon H.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Endege, Wilson O.
; APPLICANT: Ford, Donna M.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; APPLICANT: Steinmann, Kathleen E.
; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT
; FILE REFERENCE: 1532.002/200130.463
; CURRENT APPLICATION NUMBER: US/09/328.475C
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 341
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 78
; LENGTH: 1024
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(1024)
; OTHER INFORMATION: n = A,T,C or G
; US-09-328-475C-78

Query Match      21.3%; Score 142.8; DB 4; Length 1024;
Best Local Similarity 65.4%; Pred. No. 1.1e-35;
Matches 242; Conservative 0; Mismatches 122; Indels 6; Gaps 2;

QY 273 TACAGCTCTCTGAGCTGATTGGGCTATGCTTACTTGGCTCCCTTTTGTCCAGGAACCCA 332
; US-09-328-475C-77
```

```
DB 103 TACAGCTCTCTGAAATGATTGGGCTATGCGGCC---CGAGCAGTTTCAGTGATGAAGCGG 159
QY 333 CTGATGAAGAGCCCTAAAGAGAGAAACCCCACTAAAGTCGGAATCTCTACCTGATC 392
DB 160 AACAGCAACACCTTGAAGAGGGGAACCACTCAACCTCAGGATCTCTGAGCTGCTC 219
QY 393 AGAAGAGAGAGATGATCAGGCTGAGATTCAAGTGCCTGAGCTGGAAGCCGATC 452
DB 220 AGGAGGAGAGATG---AGGAGCATCTGAGGTCAAGGCCGAGCCTGAGCTGATCATA 276
QY 453 TCCAGGAGCTATGTCAGACAAAGACTCGGATGATCTGAAGTGGTACTGATGTCAGG 512
DB 277 GCCAGGAACAGGCTCACCACAGACTGGGTGTGAGTGTGAAGATGGTCTCTGATGGG 336
QY 513 GGAAGATTCTACCAAAAGCAGACACTTTAAATGCCAGAGCAGGTGAAGGGAATCAC 572
DB 337 AGATGGACCCGCAATCCAGAGGAGGTGAAGCCCTGAAGAGGTGAAGCAATCAC 396
QY 573 AGTTTAAAGGAAGATAGCTGAAACACAACTGTTTTTATATTAGATATTTTACTT 632
DB 397 AGTGTAAAGAGGACGTTGAATGATGACGCTGCTCTATGTTGGAATTTGTTCA 456
QY 633 TAAAGAGTCT 642
DB 457 TTAATAATCT 466

RESULT 43
US-09-328-475C-77/C
; Sequence 77, Application US/09328475C
; Patent No. 6476207
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jimmy
; APPLICANT: Astel, Jon H.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Endege, Wilson O.
; APPLICANT: Ford, Donna M.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; APPLICANT: Steinmann, Kathleen E.
; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT
; FILE REFERENCE: 1532.002/200130.463
; CURRENT APPLICATION NUMBER: US/09/328.475C
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 341
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 77
; LENGTH: 1024
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(1024)
; OTHER INFORMATION: n = A,T,C or G
; US-09-328-475C-77

Query Match      21.0%; Score 140.8; DB 4; Length 1024;
Best Local Similarity 64.9%; Pred. No. 4.9e-35;
Matches 240; Conservative 0; Mismatches 124; Indels 6; Gaps 2;

QY 273 TACAGCTCTCTGAGCTGATTGGGCTATGCTTACTTGGCTCCCTTTTGTCCAGGAACCCA 332
DB 518 TNCAGCTCTCTGAAATGATTGGGCTATGCGGCC---CGAGCAGTTTCAGTGATGAAGCGG 462
QY 333 CTGATGAAGAGCCCTAAAGAGAGAAACCCCACTAAAGTCGGAATCTCTACCTGATC 392
DB 461 AACAGCAACACCTTGAAGAGGGGAACCACTCAACGTGAGATCTCTGAGCTGCTC 402
QY 393 AGAAGAGAGAGATGATCAGGCTGAGATTCAGTGCCTGAGTGCCTGAGGAGCCGATC 452
DB 401 AGGAGGAGAGATG---AGGAGCATNTGAGGTCAAGGCCGAGGCTGAGCTGATCATA 345
```

QY 453 TCAGGAGCTATGTGACAGCAAGACTGGGATGGATGTGAAGTGTGCTGATGTCAAGG 512
Db 344 GCCAGGAACAGGGTCAACCCACAGACTGGTGTGAGTGTGAAGTGTGCTGATGGCAGG 285
QY 513 GGRAGATTCTACCAAGAGCAGCCTTAAATGCCAGAGCAGGTGAAGGGARATCAC 572
Db 284 AGATGGACCCGCCAAATCCAGAGAGGTGAAGACGCTGAAGAGGTGAAGCAATCAC 225
QY 573 AGGTTAAAGAGAGTAAAGCTGAAACACAACTGTTTTATATTAGATATTTTACTT 632
Db 224 AGTGTTAAAGAGGACGCTTGAATATGATGAGGCTCTCTATGTTGGAAATTTGTCA 165
QY 633 TAAGAGCTCT 642
Db 164 TTAAGATCT 155

RESULT 44

US-09-439-313-472

; Sequence 472, Application US/09439313

; Patent No. 6329505

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan Louise

; APPLICANT: Jiang Yuqui

; APPLICANT: Reed, Steven G.

; APPLICANT: Kalos, Michael

; APPLICANT: Fanger, Gary

; APPLICANT: Retter, Mark

; APPLICANT: Solk, John

; APPLICANT: Day, Craig

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND

; FILE REFERENCE: 210121.427C9

; CURRENT APPLICATION NUMBER: US/09/439,313

; CURRENT FILING DATE: 1999-11-12

; NUMBER OF SEQ ID NOS: 575

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 472

; LENGTH: 515

; TYPE: DNA

; ORGANISM: Homo sapiens

; NAME/KEY: misc feature

; LOCATION: (1)...(515)

; OTHER INFORMATION: n = A,T,C or G

US-09-439-313-472

Query Match 14.8%; Score 99.2; DB 4; Length 515;
Best Local Similarity 63.0%; Pred. No. 6.8e-22;
Matches 174; Conservative 0; Mismatches 93; Indels 9; Gaps 1;

QY 341 GAGCCTAAAGAGAGAAACCCCACTAAAGTCGGAATCTTACACCTGATCAGAAGAGA 400
Db 152 GAATCTCAGCAGAGAGAACCCCACTGACATCAGGATTTGAACCTGGCAGAGAGA 211
QY 401 GAAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAAAGCCGATCTCCAGGAG 460
Db 212 GAA-----GGAACACCTCCGATCGAAGACGTTAAAGTAGAAGGTGATTGCCAGGAA 262
QY 461 CTATGTCAGACAAAGACTGGGGATGGATGTGAAGTGTGCTGATGTCAAGGGGAAGATT 520
Db 263 ATGGATCTGGAAAGACTCGGAGTGGAGTGGAGTGGCTCTGATGTAAAGAGAGACT 322
QY 521 CTACCAAAAGCAGACACTTTTAAATGCCAGAGCAGGTGAAGGGAATCAGGTTTAA 580
Db 323 CCACCTAATCTTAAGCATGCTAAGACTAAAGAGCAGGAGATGGCGACCCATAAGTTAAA 382
QY 581 AGGAGATAGCTGAAACACCAACCAACTGTTTTAT 616
Db 383 AAGAGACAGCTGAAGCTACACATGCTGATGT 418

RESULT 45

US-09-352-616A-472

; Sequence 472, Application US/09352616A

; Patent No. 6395278

; GENERAL INFORMATION:

; APPLICANT: Dillon, Davin C.

; APPLICANT: Harlocker, Susan Louise

; APPLICANT: Jiang, Yuqui

; APPLICANT: Xu, Jiangchun

; APPLICANT: Mitcham, Jennifer Lynn

; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS

; FILE REFERENCE: 210121.427C8

; CURRENT APPLICATION NUMBER: US/09/352,616A

; CURRENT FILING DATE: 1999-07-13

; NUMBER OF SEQ ID NOS: 472

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 472

; LENGTH: 515

; TYPE: DNA

; ORGANISM: Homo sapiens

; NAME/KEY: misc feature

; LOCATION: (1)...(515)

; OTHER INFORMATION: n = A,T,C or G

US-09-352-616A-472

Query Match 14.8%; Score 99.2; DB 4; Length 515;
Best Local Similarity 63.0%; Pred. No. 6.8e-22;
Matches 174; Conservative 0; Mismatches 93; Indels 9; Gaps 1;

QY 341 GAGCCTAAAGAGAGAAACCCCACTAAAGTCGGAATCTTACACCTGATCAGAAGAGA 400
Db 152 GAATCTCAGCAGAGAGAACCCCACTGACATCAGGATTTGAACCTGGCAGAGAGA 211
QY 401 GAAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAAAGCCGATCTCCAGGAG 460
Db 212 GAA-----GGAACACCTCCGATCGAAGACGTTAAAGTAGAAGGTGATTGCCAGGAA 262
QY 461 CTATGTCAGACAAAGACTGGGGATGGATGTGAAGTGTGCTGATGTCAAGGGGAAGATT 520
Db 263 ATGGATCTGGAAAGACTCGGAGTGGAGTGGAGTGGCTCTGATGTAAAGAGAGACT 322
QY 521 CTACCAAAAGCAGACACTTTTAAATGCCAGAGCAGGTGAAGGGAATCAGGTTTAA 580
Db 323 CCACCTAATCTTAAGCATGCTAAGACTAAAGAGCAGGAGATGGCGACCCATAAGTTAAA 382
QY 581 AGGAGATAGCTGAAACACCAACCAACTGTTTTAT 616
Db 383 AAGAGACAGCTGAAGCTACACATGCTGATGT 418

Search completed: January 18, 2004, 18:48:33
Job time : 40.4394 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2004, 23:43:50 ; Search time 1008.55 Seconds
(without alignments)
16121.802 Million cell updates/sec

Title: US-10-051-835-1
Perfect score: 669
Sequence: 1 cacaacgcagccagcatt.....tttttggcagtcgatctc 669

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :

EST.*

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estnu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_esti.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_est3.*
- 12: gb_est4.*
- 13: gb_est5.*
- 14: gb_est6.*
- 15: em_estfun.*
- 16: em_estom.*
- 17: em_gss_hum.*
- 18: em_gss_inv.*
- 19: em_gss_pln.*
- 20: em_gss_vrt.*
- 21: em_gss_fun.*
- 22: em_gss_mam.*
- 23: em_gss_mus.*
- 24: em_gss_pro.*
- 25: em_gss_rtd.*
- 26: em_gss_dhg.*
- 27: em_gss_vrl.*
- 28: gb_gss1.*
- 29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	566.6	84.7	673	14	CB995986
2	558	83.4	756	10	BE903565
3	549.8	82.2	706	10	BE281431
4	545.6	81.6	780	10	BE734431

5	540	80.7	652	10	BG623213
6	538.4	80.5	592	10	AW972530
7	536.8	80.2	908	10	BG336433
8	531	79.4	782	10	BG336384
9	506.6	75.7	588	10	BE733225
10	496.8	74.3	982	10	BE900651
11	484.4	69.4	488	13	BX283754
12	445.2	66.5	533	14	T87105
13	442.6	66.2	496	9	AA993199
14	438.6	65.6	496	9	AI188535
15	438.4	65.5	543	14	N30289
16	426.6	63.8	483	9	AI149003
17	404.6	60.5	468	9	AI128458
18	404	60.4	468	9	AA039331
19	391.6	58.5	453	9	AI189715
20	377.6	56.4	532	9	AA737311
21	376.4	56.3	493	14	H53211
22	372.6	55.7	422	9	AA505794
23	372.2	55.6	512	14	W93043
24	364.6	54.5	445	10	BE302172
25	360.6	53.9	421	9	AA722699
26	357.4	53.4	437	9	AI740470
27	351.8	52.6	555	14	T87104
28	343.8	51.4	429	9	AI200158
29	334.6	50.0	404	14	H95793
30	318	47.5	349	9	AI186558
31	315.8	47.4	401	14	H53104
32	315.2	47.3	389	14	W93171
33	313.4	46.8	420	14	N42010
34	304.8	45.6	547	14	N40147
35	303.2	45.3	410	14	R99925
36	296.4	44.3	570	14	CB992682
37	296.4	44.3	570	14	CB994958
38	290.2	43.4	499	10	BG354572
39	286.4	42.8	520	14	CB961487
40	285.2	42.6	524	11	AF318372
41	280.8	42.0	503	9	AI745551
42	278.6	41.6	849	10	BG436305
43	277.8	41.5	429	9	AI189067
44	266.8	39.9	544	14	H87913
45	266.4	39.8	430	14	N27023
46	245.2	36.7	402	14	H87367
47	240.6	36.0	534	10	BG436468
48	238.4	35.6	571	14	N32547
49	234.8	35.1	347	9	AA039416
50	232.8	34.8	374	14	H96141
51	229	34.2	394	14	D59091
52	215.2	32.2	387	14	T84951
53	214.4	32.0	382	13	BX283272
54	204	30.5	414	9	AA620697
55	204	30.5	428	9	AA948244
56	204	30.5	438	9	AI954712
57	204	30.5	441	13	EX101350
58	204	30.5	486	12	BI460303
59	204	30.5	487	12	BG772292
60	204	30.5	489	10	BE881023
61	204	30.5	491	13	BU182421
62	204	30.5	499	13	BU564534
63	204	30.5	880	10	BE876234
64	204	30.5	1140	10	BE620196
65	202.4	30.3	421	9	AI631681
66	202.4	30.3	423	9	AA648117
67	201	30.0	771	13	BU929434
68	200.8	30.0	394	10	BG216048
69	200.8	30.0	485	10	BG217244
70	194.6	29.1	556	10	AW975988
71	192.2	28.7	978	13	BU532066
72	191.4	28.6	407	9	AA758618
73	188.2	28.1	487	10	BG211375
74	184.6	27.6	388	10	BG206865
75	183.4	27.4	391	10	BG205627
76	182	27.2	271	14	N41436
77	181.2	27.1	569	10	BG700165

cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLCM842 row: 1 column: 17
 High quality sequence start: 34
 High quality sequence stop: 705.
 Location/Qualifiers

FEATURES source

1. .756
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3960352"
 /tissue_type="choriocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 21"
 /note="Organ: placenta; Vector: pOTB7; Site: 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 223 a 150 c 216 g 167 t

BASE COUNT
 ORIGIN

Query Match 83.4%; Score 558; DB 10; Length 756;
 Best Local Similarity 95.8%; Pred. No. 1e-135; 5; Indels 21; Gaps 1;
 Matches 592; Conservative 0; Mismatches 5

37 TTGACACCTGCCTCAGTGTGCATGTTCACTGGGCATCTTCCCTCGACCCCTTTGCCAC 96
 Db 41 TTGACACCTGCCTCAGTGTGCATGTTCACTGGGCATCTTCCCTCGACCCCTTTGCCAC 100
 Qy 97 GTGGTACCGCTGGGAGCTGTGAGAGTGTGAGGGCAGTTCGACCGTCTGGACTCTT 156
 Db 101 GTGGTACCGCTGGGAGCTGTGAGAGTGTGAGGGCAGTTCGACCGTCTGGACTCTT 160
 Qy 157 TCTCTCTACTGAGACGCGCTATAGTTCGCGAGGCCAGTCTCTCCAGGAAGTGAATA 216
 Db 161 TCTCTCTACTGAGACGCGCTATAGTTCGCGAGGCCAGTCTCTCCAGGAAGTGAATA 220
 Qy 217 GTGAATATAGTTCGCGAGGAGATCAACATATAGGCTAGGCCAGGAGAGTTTACA 276
 Db 221 GTGAATATAGTTCGCGAGGAGATCAACATATAGGCTAGGCCAGGAGAGTTTACA 280
 Qy 277 GCCTCTGAGCTGATTGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAACCCACTGA 336
 Db 281 GCCTCTGAGCTGATTGGGCTATGCTT-----GACCCACTGA 319
 Qy 337 TGAAGAGCCTTAAAGAGAGAAACCCACTAAAGTCGGAATCTTACACCTGATCAGAA 396
 Db 320 TGAAGAGCCTTAAAGAGAGAAACCCACTAAAGTCGGAATCTTACACCTGATCAGAA 379
 Qy 397 GAGAGAGATGATCAGGGTGCAGTTCGATTCAGTTCAGTCCCTGGAAGCGATCTCCA 456
 Db 380 GAGAGAGATGATCAGGGTGCAGTTCGATTCAGTTCAGTCCCTGGAAGCGATCTCCA 439
 Qy 457 GGAGCTATGTGACACAAAGACTGGGGATGGATGTGAAGTGGTACTGTGTCAAGGGGAA 516
 Db 440 GGAGCTATGTGACACAAAGACTGGGGATGGATGTGAAGTGGTACTGTGTCAAGGGGAA 499
 Qy 517 GATTTACCAAGACGAGACATTTAAATGCCAGAGCAGGTGAGGGAATTCACAGGT 576
 Db 500 GATTTACCAAGACGAGACATTTAAATGCCAGAGCAGGTGAGGGAATTCACAGGT 559
 Qy 577 TTAAGGAAGATAGCTGAAACCAACCAAACTCTTTTATATATATATATATTTTACTTTAA 636
 Db 560 TTAAGGAAGATAGCTGAAACCAACCAAACTCTTTTATATATATATATTTTACTTTAA 619
 Qy 637 GAGTCTTAATAAATTTT 654

||||| 620 ATATCTTAATAAGTTTT 637
 BE281431 706 bp mRNA linear EST 13-JUL-2000
 G01154945F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3138515 5',
 mRNA sequence.
 BE281431
 BE281431 GI:9156450
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 706)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: rgs@bbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLCM103 row: 1 column: 12
 High quality sequence stop: 628.
 Location/Qualifiers

1. .706
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3138515"
 /tissue_type="choriocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_21"
 /note="Organ: placenta; Vector: pOTB7; Site: 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 216 a 158 c 185 g 147 t

BASE COUNT
 ORIGIN

Query Match 82.2%; Score 549.8; DB 10; Length 706;
 Best Local Similarity 95.4%; Pred. No. 1.5e-133;
 Matches 596; Conservative 0; Mismatches 7; Indels 22; Gaps 2;

Qy 36 TTGACACCTGCCTCAGTGTGCATGTTCACTGGGCATCTTCCCTTCGACCCCTTTGCCCA 95
 Db 1 TTGACACCTGCCTCAGTGTGCATGTTCACTGGGCATCTTCCCTTCGACCCCTTTGCCCA 60
 Qy 96 CGTGGTACCGCTGGGAGCTGTGAGAGTGTGAGGGCAGTTCCTCCAGCGCTCTGGACTCT 155
 Db 61 CGTGGTACCGCTGGGAGCTGTGAGAGTGTGAGGGCAGTTCCTCCAGCGCTCTGGACTCT 120
 Qy 156 TTCTCTCTACTGAGACGCGACCTTATAGGTCCGAGCCAGTCTCTCCAGGAAGTGAAT 215
 Db 121 TTCTCTCTACTGAGACGCGACCTTATAGGTCCGAGCCAGTCTCTCCAGGAAGTGAAT 180
 Qy 216 AGTGAATATAGTTCGCGAGGAGATCAACATATAGGCTAGGCCAGGAGAGTTTAC 275
 Db 181 AGTGAATATAGTTCGCGAGGAGATCAACATATAGGCTAGGCCAGGAGAGTTTAC 240
 Qy 276 AGCTCTCTGAGCTGATTGGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAACCCACTG 335
 Db 241 AGCTCTCTGAGCTGATTGGGGCTATGCTT-----GACCCACTG 279

```

336 ATGAAGAGCCTAAAGAGAGAAACCCACCACTAAAAAGTCGGAATCCTACACCTGATCAGA 395
Db      |
280 ATGAAGAGCCTAAAGAGAGAAACCCACCACTAAAAAGTCGGAATCCTACACCTGATCAGA 339
Qy      |
396 AGAGAGAGATGATCAGGCTGACGTGAGATTCAGTTCAGTGCCTGACCTGGAGCGGATCTCC 455
Db      |
340 AGAGAGAGATGATCAGGCTGACGTGAGATTCAGTTCAGTGCCTGACCTGGAGCGGATCTCC 399
Qy      |
456 AGAGAGATGATCAGAGCAAAAGACTGGGGATGGATGTGAAGTGGTACTGATGTCAGAGGGGA 515
Db      |
400 AGAGAGATGATCAGAGCAAAAGACTGGGGATGGATGTGAAGTGGTACTGATGTCAGAGGGGA 459
Qy      |
516 AGATTCACCAAAAGCAGACACTTTAAATGCCAGAGCAGGTGAAGGGAATCAGAG 575
Db      |
460 AGATTCACCAAAAGCAGACACTTTAAATGCCAGAGCAGGTGAAGGGAATCAGAG 519
Qy      |
576 TTTA-AAGGAGATAGCTGAAACACACAACTGTTTTATATTAGATATTTACTTTA 634
Db      |
520 TTACAGGAGATAGCTGAAACACACAACTGTTTTATATTAGATATTTACTTTA 579
Qy      |
635 AAGAGCTCTTAATAATTTTGGCAT 659
Db      |
580 AATAATCTTAATAAGTTTAAGCTT 604

RESULT 4
BE734431
LOCUS      780 bp      mRNA      linear      EST 15-SEP-2000
DEFINITION 601566206F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3841272 5',
mRNA sequence.
ACCESSION BE734431
VERSION    BE734431.1 GI:10148423
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1. (bases 1 to 780)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished
            Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-x@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LCM532 row: k column: 01
            High quality sequence stop: 652.
FEATURES   Location/Qualifiers
            1..780
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:3841272"
                /tissue_type="choriocarcinoma"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NIH_MGC_21"
            note="Organ: placenta; Vector: pOTB7; Site: 1; XhoI;
            Site 2: EcoRI; cDNA made by oligo-dT priming.
            Directionally cloned into EcoRI/XhoI sites using the
            following 5' adaptor: GGCACGAG(G). Size-selected >500bp
            for average insert size 1.8kb. Library constructed by
            Ling Hong in the laboratory of Gerald M. Rubin (University
            of California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 247 a

Query Match 81.6%; Score 545.6; DB 10; Length 780;

```

Best Local Similarity 95.9%; Pred. No. 1.9e-132;
Matches 579; Conservative 0; Mismatches 4; Indels 21; Gaps 1;
Qy      |
51 AGTGTGATGTTTCACTGGGATCTTCCCTTCGAGACCCCTTTGCCCCACAGTGTGACCGCTGG 110
Db      |
1 AGTGTGATGTTTCACTGGGATCTTCCCTTCGAGACCCCTTTGCCCCACAGTGTGACCGCTGG 60
Qy      |
111 GGACTGTGAGAGTGTGAGGGGACGTTCCAGCGGTCTGGACTCTTCTCTCTCTACTAGAG 170
Db      |
61 GGACTGTGAGAGTGTGAGGGGACGTTCCAGCGGTCTGGACTCTTCTCTCTCTACTAGAG 120
Qy      |
171 ACGCAGCCTATPAGTCCGCGCAGGCGAGTCTCCAGGAACTGAAATAGTGAATATAGATT 230
Db      |
121 ACGCAGCCTATPAGTCCGCGCAGGCGAGTCTCCAGGAACTGAAATAGTGAATATAGATT 180
Qy      |
231 GGCAGGAGAGATCAACATATAGGCTAGGCTAGGCCAGAGAAAGTTTACAGCCTCTCTGAGCTGA 290
Db      |
181 GGCAGGAGAGATCAACATATAGGCTAGGCTAGGCCAGAGAAAGTTTACAGCCTCTCTGAGCTGA 240
Qy      |
291 TTGGGGCTATCTTACTTGGCTCCCTTTGTCCAGGAACTCCCTGATGAAGAGCCTAAAG 350
Db      |
241 TTGGGGCTATCTTACTTGGCTCCCTTTGTCCAGGAACTCCCTGATGAAGAGCCTAAAG 279
Qy      |
351 AAGGAGAAACCCACCTAAAGTCGGAATCCTACACCTGATCAGAGAGAGAGATGATC 410
Db      |
280 AAGGAGAAACCCACCTAAAGTCGGAATCCTACACCTGATCAGAGAGAGAGATGATC 339
Qy      |
411 AGGGTGCACTGAGATTCAGTGCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTCAGA 470
Db      |
340 AGGGTGCACTGAGATTCAGTGCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTCAGA 399
Qy      |
471 CAAGACTGGGGATGGATGTGAAGTGGTACTGATGTCAAGGGGAGAGATCTTACCAAAAG 530
Db      |
400 CAAGACTGGGGATGGATGTGAAGTGGTACTGATGTCAAGGGGAGAGATCTTACCAAAAG 459
Qy      |
531 CAGAGCACTTTAAATGCCAGACAGTGAAGGGAATCACAGGTTTAAAGGAAGATAA 590
Db      |
460 CAGAGCACTTTAAATGCCAGACAGTGAAGGGAATCACAGGTTTAAAGGAAGATAA 519
Qy      |
591 GCTGAACACACAACTGTTTTATATTAGATATTTTACCTTTAAAGAGTCTTAATAAT 650
Db      |
520 GCTGAACACACAACTGTTTTATATTAGATATTTTACCTTTAAAGAGTCTTAATAAT 579
Qy      |
651 TTTT 654
Db      |
580 TTTT 583

RESULT 5
BE623213
LOCUS      652 bp      mRNA      linear      EST 18-APR-2001
DEFINITION 602648125F1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4769551 5',
mRNA sequence.
ACCESSION BE623213
VERSION    BE623213.1 GI:13674584
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1. (bases 1 to 652)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished
            Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-x@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
```

Plate: L1CM1635 row: a column: 08
 High quality sequence stop: 627.
 Location/Qualifiers
 1. .652
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4769551"
 /lab_host="DH10B (T1 phage-resistant)"/
 /clone_lib="NIH_MGC_79"
 /notes="Organ: placenta; Vector: pDNR-LIB (Clontech); Site 1: SfilI (ggccctatggcc); Site 2: SfilI (ggccattatggcc) ; 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGCGCATTTAGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGAGGCGCCACATG-dT(30)BN-3' (where B = A, C, G and N = A, C, G, or T). Average insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."
 BASE COUNT 192 a 144 c 165 g 151 t
 ORIGIN
 Query Match 80.7%; Score 540; DB 10; Length 652;
 Best Local Similarity 95.4%; Pred. No. 5.4e-131;
 Matches 607; Conservative 0; Mismatches 5; Indels 24; Gaps 4;
 QY 22 AGTGTGCATGTTCTCTTGACACCTGCCCTCAGTGTGCATGTTCACTGGGCATCTTCCCTTC 81
 Db 2 AGTGTGCATGTTCTCTTGACACCTGCCCTCAGTGTGCATGTTCACTGGGCATCTTCCCTTC 61
 QY 82 GACCCCTTTCGCCCGCTGCACCGCTGGGAGCTGTGAGAGTGTGAGGGGCAGCTTCCA 141
 Db 62 GACCCCTTTCGCCCGCTGCACCGCTGGGAGCTGTGAGAGTGTGAGGGGCAGCTTCCA 121
 QY 142 GCCGCTCGCATCTTTCTCTCTACTGAGACCGACCTTATAGTTCGCGAGCGCATGCTTC 201
 Db 122 GCCGCTCGCATCTTTCTCTCTACTGAGACCGACCTTATAGTTCGCGAGCGCATGCTTC 181
 QY 202 CCAGGAATCGAAATAGTCAAAATATGAGTTGGCGAGGAGAGATCAACATATAGCGCTAGGCC 261
 Db 182 CCAGGAATCGAAATAGTCAAAATATGAGTTGGCGAGGAGAGATCAACATATAGCGCTAGGCC 241
 QY 262 AAGAAAGAGTTTACAGCCCTCTGACCTGATTGGGCTATGCTTACTGTGCTCCCTTTGTC 321
 Db 242 AAGAAAGAGTTTACAGCCCTCTGACCTGATTGGGCTATGCTT----- 284
 QY 322 CCAGGAACCCACTGATGAGACGCTTAAAGAGAGAAACCACCCATCTAAAGTCGGAATCC 381
 Db 285 -----GAAACCACTGATGAGACCTTAAAGAGAGAAACCACCCATCTAAAGTCGGAATCC 340
 QY 382 TACACCTGATCAGAGAGAGAA- GATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACC 440
 Db 341 TACACCTGATCAGAGAGAGAAACGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACC 400
 QY 441 TCGAAGCCGATCTCCAGGAGCTA-TGTTCAGACAAAGACCTGGGCGATGGATGTCAAGTGGT 499
 Db 401 TCGAAGCCGATCTCCAGGAGCTATTGTTCAGACAAAGACCTGGGCGATGGATGTGAAGTGGT 460
 QY 500 ACTGATGTCAAGGGGAAGATTCTACAAAGCAGAGCACTTTAAATGCCAGAGCAGGT 559
 Db 461 ACTGATGTCAAGGGGAAGATTCTACAAAGCAGAGCACTTTAAATGCCAGAGCAGGT 520
 QY 560 GAAGGGAATCACAGGTTTAAAGGAGAGATAAGCTGAAACCAACAACTGTTTTTATA-T 618
 Db 521 GAAGGGAATCACAGGTTTAAAGGAGAGATAAGCTGAAACCAACAACTGTTTTTATACT 580
 QY 619 TAGATATTTTACTTTAAAGAGCTTTAAATATTTT 654
 Db 581 TAGATATTTTACTTTAAATATCTTAATTAAGTTTT 616

RESULT 6
AW972530/c
LOCUS AW972530 592 bp mRNA linear EST 01-JUN-2000
DEFINITION ESI384621 WAGE resequences, MAGL Homo sapiens cDNA, mRNA sequence.
ACCESSION AW972530
VERSION AW972530.1 GI:8162376
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 592);
AUTHORS Hegde,P., Q.R.; Abernathy,K., Dharap,S., Gasparid,R., Gay,C., Holt,
,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
Quackenbush,J.
TITLE Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
Plate: j07
Seq primer: Forward.
Location/Qualifiers
source 1..592
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone_lib="WAGE resequences, MAGL"
/note="Vector: pluescriptSKm"
BASE COUNT 138 a 152 C 122 G 179 t 1 others
ORIGIN
Query Match 80.5%; Score 538.4; DB 10; Length 592;
Best Local Similarity 95.3%; Pred. No. 1.4e-130;
Matches 573; Conservative 0; Mismatches 7; Indels 21; Gaps 1;
QY 54 GTGCATGTTCACTGGCGCATCTTCCTTCCAGCCCCTTTGCCACGPGTGACCGCTGGGA 113
DB 591 GTGCATGTTCAATTGGGCATCTTCCATTCGACCCTTTGCCACGTGGTAGCCTGGGA 532
QY 114 CTGTGAGAGTGNGGGGACGTTCAGCGCTCTGGACTCTTCTCTCCCTACTCAGACG 173
DB 531 NCTGTGAGAGTGAGGGGACGTTCCAGCGCTCTGGACTCTTCTCTCTCTACTGACG 472
QY 174 CAGCCTATAGTCCCGAGCCAGTCCTCCCCAGAACTCAAATAGTGAATATGAGTTGCC 233
DB 471 CAGCCTATAGTCCCGAGCCAGTCCTCCCCAGAACTCAAATAGTGAATATGAGTTGCC 412
QY 234 GAGGAAGATCAACATATAGGCTAGGCCAAGAAGAAGTTTACAGCTCTCTGAGCTGATTG 293
DB 411 GAGGAAGATCAACATATAGGCTAGGCCAAGAAGAAGTTTACAGCTCTCTGAGCTGATTG 352
QY 294 GGGCTATGCTTAAGTGGCTCCCTTTGTCCCAGAACCCCACTGATGAAGAGCCTAAGAAG 353
DB 351 GGGCTATGCTT-----GAAACCCACTGATGAAGAGCCTAAGAAG 313
QY 354 AGAAACCCACTAAAGTGGGAATCTTACACCTGATCAGAAAGAGAGAAGATGATCAGG 413
DB 312 AGAAACCCACTAAAGTGGGAATCTTACACCTGATCAGAAAGAGAGAAGATGATCAGG 253
QY 414 GTGCAGCTGAGATTCAAGTGCTGACCTGGAAAGCCGATCTCCAGGAGCTATGTFCAGACA 473
DB 252 GTGCAGCTGAGATTCAAGTGCTGACCTGGAAAGCCGATCTCCAGGAGCTATGTFCAGACA 193
QY 474 AGACTGGGATGGATGTGAAGTGGTGTCTGAGTGTCTAAGGGGGAAGATTTCTACCAAAGACG 533
DB 192 AGACTGGGATGGATGTGAAGTGGTGTCTGAGTGTCTAAGGGGGAAGATTTCTACCAAAGACG 133
QY 534 AGCACTTTAAAATGCCAGAACGAGGTGAAGGGGAATCA CAGGTTTTAAGGAAGATAGCT 593

of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 240 a 155 c 218 g 169 t

ORIGIN

Query Match 79.4%; Score 531; DB 10; Length 782;
Best Local Similarity 94.6%; Pred. No. 1.3e-128;
Matches 591; Conservative 0; Mismatches 10; Indels 24; Gaps 3;

QY 31 GTTCTTGGACACCTGCTCAGTGTGCATGTTCACTGGGCATCTTCCCTCGACCCCTTT 90
Db 2 GTTCTTGGACACCTGCTCAGTGTGCATGTTCACTGGGCATCTTCCCTCGACCCCTTT 61
QY 91 GCCACATGTTGACCGCTGGGGAGCTGTGAGAGTGTGAGGGGACAGTTCAGCGCTGG 150
Db 62 GCCACATGTTGACCGCTGGGGAGCTGTGAGAGTGTGAGGGGACAGTTCAGCGCTGG 121
QY 151 ACTCTTTCTCTCTACTGACCGCAGCTATAGTTCGGAGGCGAGTCTCCAGGAAGT 210
Db 122 ACTCTTTCTCTCTACTGACCGCAGCTATAGTTCGGAGGCGAGTCTCCAGGAAGT 181
QY 211 GAAATAGTGAATATGAGTTGGGAGGAGGATCAACATATAGGCTAGGCCAGAGAG 270
Db 182 GAAATAGTGAATATGAGTTGGGAGGAGGATCAACATATAGGCTAGGCCAGAGAG 241
QY 271 TTACAGCCTCTGAGCTGATGGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAAC 330
Db 242 TTACAGCCTCTGAGCTGATGGGGCTATGCTT-----GAAC 280
QY 331 CACTGATGAAGAGCTAAAGAGAGAACACCCACCTAAAGTCGGAATCTACACTGA 390
Db 281 CACTGATGAAGAGCTAAAGAGAGAACACCCACCTAAAGTCGGAATCTACACTGA 340
QY 391 TCAGAGAGAGAGATGATCAGGGTGAGCTGAGATTCAAGTGCCTGACCTGGAAGCCGA 450
Db 341 TCAGAGAGAGAGATGATCAGGGTGAGCTGAGATTCAAGTGCCTGACCTGGAAGCCGA 400
QY 451 TCTCCAGAGCTATGTCAGACAAAGACTGGGGATGAGTGAAGTGGTACTGATGCAA 510
Db 401 TCTCCAGAGCTATGTCAGACAAAGACTGGGGATGAGTGAAGTGGTACTGATGCAA 460
QY 511 GGGAGAGTCTACCAAAAGCAGAGCATTAAATGCCAGAGCAGGTGAAGGGAATC 570
Db 461 GGGAGAGTCTACCAAAAGCAGAGCATTAAATGCCAGAGCAGGTGAAGGGAATC 520
QY 571 ACAGTTTAAAGGAGATAGCTGAAACACACAACTGTTTATATTAGATATTAC 630
Db 521 ACAGTTTAAAGGAGATAGCTGAAACACAACTGTTTATATTAGATATTAC 578
QY 631 TTTAAGA-GTCTTAAATAATTTT 654
Db 579 TTTAATAATCTTAAATAATTTT 603

RESULT 9
BE733225 588 bp mRNA linear EST 15-SEP-2000
LOCUS 601568462P1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3843231 5',
DEFINITION mRNA sequence.
ACCESSION BE733225
VERSION BE733225.1 GI:10147217
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 588)
NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L2CM537 row: 1 column: 16
High quality sequence stop: 588.

FEATURES
Location/Qualifiers
1..588
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3843231"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_21"
/note="Organ: placenta; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >50bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 178 a 124 c 152 g 134 t

ORIGIN

Query Match 75.7%; Score 506.6; DB 10; Length 588;
Best Local Similarity 95.3%; Pred. No. 3.2e-122;
Matches 573; Conservative 0; Mismatches 4; Indels 24; Gaps 4;

QY 54 GTGCATGTTCACTGGGCATCTTCCCTTCGACCCCTTTGCCACGCTGGTGGCGGA 113
Db 1 GTGCATGTTCACTGGGCATCTTCCCTTCGACCCC-TTGCCACGCTGGTGGCGGA 59
QY 114 GCTGTGAGAGTGTGAGGGCACGTTCCAGCCGTCTGGACTCTTTCTCTCTACTGAGACG 173
Db 60 GCTGTGAGAGTGTGAGGGCACGTTCCAGCCGTCTGGACTCTTTCTCTCTACTGAGACG 119
QY 174 CAGCCTATAGTTCGCGAGCCAGTCCCTCCAGGAACCTGAATATGTAATGATGTTGGC 233
Db 120 CAGCCTATAGTTCGCGAGCCAGTCCCTCCAGGAACCTGAATATGTAATGATGAG-TGGC 178
QY 234 GAGGAAGATCAACATATATAGGCTTAGGGCCAGAGAAAGTTTACAGCCTCCTGAGCTGATTG 293
Db 179 GAGGAAGATCAACATATATAGGCTTAGGGCCAGAGAAAGTTTACAGCCTCCTGAGCTGA-TG 237
QY 294 GGGCTATGCTTACTGGCTCCCTTTGTCCAGGAACCCACTGATGAAGAGCTTAAAGAG 353
Db 238 GGGCTATGCTT-----GAAACCCACTGATGAAGAGCTTAAAGAG 276
QY 354 AGAAACCCACCCACTAAAGTGGGAATCTTACACCTGATCAGAGAGAGAGATGATCAGG 413
Db 277 AGAAACCCACCCACTAAAGTGGGAATCTTACACCTGATCAGAGAGAGAGATGATCAGG 336
QY 414 GTGCAGTGAAGATTCAGTGCCTGACCTGGAGCGGATCTCCAGAGCTATGTCAGACAA 473
Db 337 GTGCAGTGAAGATTCAGTGCCTGACCTGGAGCGGATCTCCAGAGCTATGTCAGACAA 396
QY 474 AGACTGGGGATGGATGTCAAGTGGTACTGTATGTCAAGGGGAAGATTTACCAAAAGCAG 533
Db 397 AGACTGGGGATGGATGTCAAGTGGTACTGTATGTCAAGGGGAAGATTTACCAAAAGCAG 456
QY 534 AGCACTTTAAATGCCAGAGCAGGTGAAGGGAATCAACAGTTTAAAGGAAGATAGCT 593
Db 457 AGCACTTTAAATGCCAGAGCAGGTGAAGGGAATCAACAGTTTAAAGGAAGATAGCT 516
QY 594 GAAACACACAACTGTTTATATTAGATATTTTACTTTTAAAGAGTCTTAAATAATTTT 653
Db 517 GAAACACACAACTGTTTATATTAGATATTTTACTTTTAAATAATTTTAAATAATTTT 576
QY 654 T 654

Db 577 T 577

RESULT 10
BE900651
LOCUS
DEFINITION
601673578f1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3956491 5',
mRNA sequence.
ACCESSION
BE900651
VERSION
EST.
KEYWORDS
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 982)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LCMW832 row: k column: 20
High quality sequence start: 3
High quality sequence stop: 637.
Location/Qualifiers
1..982
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3956491"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_21"
/note="Organ: placenta; Vector: pOTB7; Site: 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dt priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 312 a 223 c 276 g 171 t

Query Match 74.3%; Score 496.8; DB 10; Length 982;
Best Local Similarity 94.8%; Pred. NO. 1.4e-119; Indels 24; Gaps 4;
Matches 565; Conservative 0; Mismatches 7;
QY 64 ACTGGGCATCTTCCCTTCGACCCCTTTGCCACGTTGGTGACCGCTGGGAGCTGTGAGAG 123
DB 1 ACTGGGCATCTTCCCTTCGACCCCTTTGCCACGTTGGTGACCGCTGGGAGCTGTGAGAG 59
QY 124 TGTGAGGGGACGTTCCAGCGCTCGACTCTTCTCTCTACTGAGCGCGCTATAG 183
DB 60 TGTGAGGGGACGTTCCAGCGCTCGACTCTTCTCTCTACTGAGCGCGCTATAG 119
QY 184 GTCCGAGGGCAGTCTCTCCAGGAAGTGAATAGTGAATATGATGTCGCGAGGAATC 243
DB 120 GTCCGAGGGCAGTCTCTCCAGGAAGTGAATAGTGAATATGATGTCGCGAGGAATC 178
QY 244 AACATATAGGCTTAGGCAAGAGAGATTACAGCTCTGAGCTATTTGGGGCTATGCT 303
DB 179 AACATATAGGCTTAGGCAAGAGAGATTACAGCTCTGAGCTATTTGGGGCTATGCT 238
QY 304 TACTGGCTCCCTTTGTGCCAGGAACCCACTGATGAAGAGCCCTTAAGAGAGAAACACC 363
DB 239 T-----GACCCACTGATGAAGAGCCCTTAAGAGAGAAACACC 277

QY 364 CACTAAAGTCGGAAATCCTACCTGATCAGAAAGAGAAAGATGATCAGGGTGCAGCTGA 423
DB 278 CACTAAAGTCGGAAATCCTACCTGATCAGAAAGAGAAAGATGATCAGGGTGCAGCTGA 337
QY 424 GATTCAAGTGCCTGACCTGGAGCCGATCTCCAGGAGCTATGTTCAGACAAAGACTGGGGA 483
DB 338 GATTCAAGTGCCTGACCTGGAGCCGATCTCCAGGAGCTATGTTCAGACAAAGACTGGGGA 397
QY 484 TGGATGTGAAGTGGTACTGATGTCAAGGGGAAGATTCTACCAAAAGCAGAGCACTTTAA 543
DB 398 TGGATGTGAAGTGGTACTGATGTCAAGGGGAAGATTCTACCAAAAGCAGAGCACTTTAA 457
QY 544 AATGCAGAACAGCTGGAAGGGAATCAAGGTTAAAGGAAGATAGTCAAAACACAC 603
DB 458 AATG-CAGAAGCAGCTGGAAGGGAATCAAGGTTAAAGGAAGATAGTCAAAACACAC 516
QY 604 AAACCTGTTTTATATTAGATATTTTACTTAAAGAGTCTTAATAAATTTTGGCAT 659
DB 517 AAACCTGTTTTATATTAGATATTTTACTTAAATATCTTTAAAGTTTAAAGCTT 572

RESULT 11
BX283754
LOCUS
DEFINITION
498 bp mRNA linear EST 05-MAR-2003
IMAGE:3960352, mRNA sequence.
ACCESSION
BX283754
VERSION
EST.
KEYWORDS
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 498)
Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
Radelof, U., Schneider, D. and Korn, B.
Human Unigeneset - RZPD3
Unpublished
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 380, D-69120 Heidelberg, Germany
RZPD; IMAGE:3960352
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
<http://www.rzpd.de/CiOneCards/cgi-bin/showlib.pl.cgi/response?libNo=972> Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
PCMV-M13, Primer sequence: CGTTGTAAACGACGGCCAGT.

FEATURES
Location/Qualifiers
1..498
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3960352"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_21"
/note="Organ: placenta; Vector: pOTB7; Site: 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dt priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 135 a 117 c 139 g 107 t


```

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 496)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 634 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 474.
FEATURES
    source
        1..496
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:1624334"
            /dev_stage="8-9 weeks"
            /lab_host="DH10B"
            /clone_lib="Soares total_fetus_Nb2HF8_9w"
            /note="Vector: p773D-Pac (Pharmacia) with a modified
            polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
            was prepared from mRNA obtained from pooled 8-9 week
            (total) fetus material with a Not I - oligo(dT) primer [5'
            TGTTACCAATCTGAAGTGGGAGCGCGCGCTTAATTTTTTTTTTTT 3'].
            Double-stranded cDNA was ligated to Eco RI adaptors
            (Pharmacia), digested with Not I and cloned into the Not I
            and Eco RI sites of the modified p773 vector. Library
            went through one round of normalization, and was
            constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 112 a 119 c 96 g 169 t
ORIGIN
    Query Match 66.2%; Score 442.6; DB 9; Length 496;
    Best Local Similarity 95.0%; Pred. No. 2.1e-105;
    Matches 476; Conservative 0; Mismatches 4; Indels 21; Gaps 1;
QY 154 CTTTCTCTCTACTGAGCGGAGCTATAGTTCGCGAGCGCAGTCTCCCGAGGAAGTAA 213
DB 496 CTTTCTCTCTCTACTGAGCGGAGCTATAGTTCGCGAGCGCAGTCTCCCGAGGAAGTAA 437
QY 214 ATAGTGAATATAGTTCGCGAGGAGATCAACATATATAGGCTAGGCCAAGAAAGTTT 273
DB 436 ATAGTGAATATAGTTCGCGAGGAGATCAACATATATAGGCTAGGCCAAGAAAGTTT 377
QY 274 ACAGCTCTCTGAGCTGATTGGGGCTATGCTTACTGCTCCCTTTGTCGAGGAACCCAC 333
DB 376 ACAGCTCTCTGAGCTGATTGGGGCTATGCTTACTGCTCCCTTTGTCGAGGAACCCAC 338
QY 334 TGATGAAGAGCCTTAAAGAGAGAAACCCACCTAAAGTGGAAATCTTACACCTGATCA 393
DB 337 TGATGAAGAGCCTTAAAGAGAGAAACCCACCTAAAGTGGAAATCTTACACCTGATCA 278
QY 394 GAAGAGAGAGATGATCAGGTGCGAGTTCAGATTCAGTTCGCTGACCTGGAGCCGATCT 453
DB 277 GAAGAGAGAGATGATCAGGTGCGAGTTCAGATTCAGTTCGCTGACCTGGAGCCGATCT 218
QY 454 CCAGGAGCTATGTGCAGCAAAAGACTGGGGATGGATGTGAAGGTGGTACTGATGTCAAGGG 513
DB 217 CCAGGAGCTATGTGCAGCAAAAGACTGGGGATGGATGTGAAGGTGGTACTGATGTCAAGGG 158
QY 514 GAAGATTCTACAAAAGCAGAGACACTTTAAATGCCAAGAGCAGGTGAAGGAATCACA 573
DB 157 GAAGATTCTACAAAAGCAGAGACACTTTAAATGCCAAGAGCAGGTGAAGGAATCACA 98
QY 574 GGTTTAAAGAGATAGCTGAAACCAACAACTGTTTTTATATAGATATTTTACTTTT 633

```

```

97 GGTTTAAAGAGATAGCTGAAACCAACAACTGTTTTTATATATAGATATTTTACTTT 38
634 AAAGAGCTCTTAATAAATTTT 654
37 AAAATATCTTTAATAAAGTTT 17
RESULT 14
LOCUS A1188535/c
DEFINITION 496 bp mRNA linear EST 28-OCT-1998
cd09e11.x1 Soares placenta_8to9weeks_2NBHP8to9w Homo sapiens cDNA
clone IMAGE:1723244_3' similar to SW:GGE2_HUMAN Q13066 GAGE-2
PROTEIN. [1] ; mRNA sequence.
ACCESSION A1188535
VERSION A1188535.1 GI:3739744
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 601 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 473.
FEATURES
    source
        1..496
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:1723244"
            /dev_stage="two placentae: one from 8 weeks and another
            from 9 weeks post conception"
            /lab_host="DH10B (ampicillin resistant)"
            /clone_lib="Soares placenta_8to9weeks_2NBHP8to9w"
            /note="Organ: placenta; Vector: p773D (Pharmacia) with a
            modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
            strand cDNA was primed with a Not I - oligo(dT) primer [5'
            TGTTACCAATCTGAAGTGGGAGCGCGCGATTTTTTTTTTTT 3'],
            double-stranded cDNA was size selected, ligated to Eco RI
            adaptors (Pharmacia), digested with Not I and cloned into
            the Not I and Eco RI sites of a modified p773 vector
            (Pharmacia). Library constructed by Bento Soares and
            M. Fatima Bonaldo."
BASE COUNT 109 a 119 c 95 g 173 t
ORIGIN
    Query Match 65.6%; Score 438.6; DB 9; Length 496;
    Best Local Similarity 95.0%; Pred. No. 2.3e-104;
    Matches 472; Conservative 0; Mismatches 4; Indels 21; Gaps 1;
QY 158 CTCTCTCTACTGAGACGCGAGCTATAGTTCGCGAGCGCAGTCTCCCGAGGAAGTAAATAG 217
DB 496 CTCTCTCTACTGAGACGCGAGCTATAGTTCGCGAGCGCAGTCTCCCGAGGAAGTAAATAG 437
QY 218 TGAATATATGTTGGCGAGGAAGATCAACATATAGGCTAGGCCAAGAGAGAGTTTACAG 277
DB 436 TGAATATATGTTGGCGAGGAAGATCAACATATAGGCTAGGCCAAGAGAGAGTTTACAG 377
QY 278 CCTCTGAGCTGATTGGGGCTATGCTTACTGCTCCCTTTGTCGAGGAACCCACTGAT 337
DB 376 CCTCTGAGCTGATTGGGGCTATGCTTACTGCTCCCTTTGTCGAGGAACCCACTGAT 338
QY 338 GAAGAGCCTTAAAGAGAGAAACCCACCTAAAGTGGAAATCTTACACCTGATCTCAGAG 397
DB 337 GAAGAGCCTTAAAGAGAGAAACCCACCTAAAGTGGAAATCTTACACCTGATCTCAGAG 278

```


/db_xref="taxon:9606"
 /clone="IMAGE:1715402"
 /dev_stage="two placentae: one from 8 weeks and another
 from 9 weeks post conception"
 /lab_host="PH108 (ampicillin resistant)"
 /clone_lib="Soares_placenta_25weeks_2MBHPto9W"
 /notes="Organ: placenta; Vector: p773D (Pharmacia) with a
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5'
 TGTTCACATCTGAAGTGGGAGCGCGCGGATTTTTTTTTTTT 3']
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified p773 vector
 (Pharmacia). Library constructed by Bento Soares and
 M. Fatima Bonaldo."

BASE COUNT 105 a 119 c 90 g 169 t

Query Match 63.8%; Score 426.6; DB 9; Length 483;
 Best Local Similarity 94.8%; Pred. No. 3.3e-101;
 Matches 460; Conservative 0; Mismatches 4; Indels 21; Gaps 1;

QY 170 GAGCGAGCTATAGTCCCGAGCCAGTCTCCAGGAACTGAATAGTGAATATGAT 229
 Db 483 GAGCGAGCTATAGTCCCGAGCCAGTCTCCAGGAACTGAATAGTGAATATGAT 424
 QY 230 TGGCGAGGAGATCAACATATAGGCTAGGCCCAAGAGAGTTTACAGCTCTCTGAGCTG 289
 Db 423 TGGCGAGGAGATCAACATATAGGCTAGGCCCAAGAGAGTTTACAGCTCTCTGAGCTG 364
 QY 290 ATTGGGGCTATGCTTACTGGCTCCCTTTGTCGAGGACCCACTGATGAAGCCCTAAA 349
 Db 363 ATTGGGGCTATGCTTACTGGCTCCCTTTGTCGAGGACCCACTGATGAAGCCCTAAA 325
 QY 350 GAAGAGAAACCCACCACTAAAAGTGGAACTCTACACCTGATCAGAGAGAGAGATGAT 409
 Db 324 GAAGAGAAACCCACCACTAAAAGTGGAACTCTACACCTGATCAGAGAGAGAGATGAT 265
 QY 410 CAGGGTGACGTGAGATTCAAGTGGCTGACCTGGAGCGGATCTCCAGGAGCTATGTCAG 469
 Db 264 CAGGGTGACGTGAGATTCAAGTGGCTGACCTGGAGCGGATCTCCAGGAGCTATGTCAG 205
 QY 470 ACAAGACTGGGATGGATGTGAAGTGGTACTGTGTCAAGGGGAAGATTCTACCAAAA 529
 Db 204 ACAAGACTGGGATGGATGTGAAGTGGTACTGTGTCAAGGGGAAGATTCTACCAAAA 145
 QY 530 GCAGAGCATTTAAATGCGAGAGCAGGTGAGGGAATCAGAGTTTAAAGGAAGATA 589
 Db 144 GCAGAGCATTTAAATGCGAGAGCAGGTGAGGGAATCAGAGTTTAAAGGAAGATA 85
 QY 590 AGCTGAACCAACAACTGTTTTATATTAGATATTTTACTTTTAAAGAGTCTTTAATAAA 649
 Db 84 AGCTGAACCAACAACTGTTTTATATTAGATATTTTACTTTTAAAGAGTCTTTAATAAA 25
 QY 650 TTTT 654
 Db 24 GTTTT 20

RESULT 17
 A1128458/c
 LOCUS
 DEFINITION
 IMAGE:1711966 3', similar to SW:GG2_HUMAN Q13066 GAGE-2 PROTEIN. [1
]; mRNA sequence.

ACCESSION A1128458
 VERSION A1128458.1 GI:3596972
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 468)

AUTHORS
 TITLE
 JOURNAL
 COMMENT

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 This clone is available royalty-free through LLNL / contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 596 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 444.

FEATURES
 source

Location/Qualifiers
 1..468
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1711966"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Soares_pregnant_uterus_NbHPU"
 /note="Organ: uterus; Vector: p773-Pac; Site 1: Not I;
 Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
 oligo(dT) primer [5'
 AACTGGAAGATTCGCGCGCCCTTTTTTTTTTTT 3']
 double-stranded cDNA was ligated to Eco RI adapters
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified p773 vector. Library
 went through one round of normalization. Library
 constructed by M. Fatima Bonaldo."

BASE COUNT 102 a 111 c 83 g 172 t

Query Match 60.5%; Score 404.6; DB 9; Length 468;
 Best Local Similarity 94.6%; Pred. No. 2e-95;
 Matches 438; Conservative 0; Mismatches 4; Indels 21; Gaps 1;

QY 192 GCCATGCTCCCGAGAACTGAATAGTGAATATGATGAGTTGCGGAGGAGATCAACATATA 251
 Db 468 GCCATGCTCCCGAGAACTGAATAGTGAATATGATGAGTTGCGGAGGAGATCAACATATA 409
 QY 252 GGCCTAGGCGCAAGAGAGTTTACAGGCTCTCTGAGCTGATTTGGGGCTATGCTTACTGGCT 311
 Db 408 GGCCTAGGCGCAAGAGAGTTTACAGGCTCTCTGAGCTGATTTGGGGCTATGCTTACTGGCT 356
 QY 312 CCCCTTTGTCGAGAACCCACTGATGAGAGCCCTTAAGAGAGCAACCCCACTAAAA 371
 Db 355 -----GAACCCACTGATGAGAGCCCTTAAGAGAGCAACCCCACTAAAA 310
 QY 372 GTCGGAATCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGCGAGCTGAGATTCAAG 431
 Db 309 GTCGGAATCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGCGAGCTGAGATTCAAG 250
 QY 432 TGCCCTGACCTGGAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGGATGGATG 491
 Db 249 TGCCCTGACCTGGAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGGATGGATG 190
 QY 492 AAGGTGGTACTGATGTCAAGGGGAAGATTCTACAAAGCAGAGCACTTTTAAATGCCAG 551
 Db 189 AAGGTGGTACTGATGTCAAGGGGAAGATTCTACAAAGCAGAGCACTTTTAAATGCCAG 130
 QY 552 AAGCAGGTGAAGGGAATCAACAGCTTTAAAGGAGATTAAGCTGAACCAACCAACTGTT 611
 Db 129 AAGCAGGTGAAGGGAATCAACAGCTTTAAAGGAGATTAAGCTGAACCAACCAACTGTT 70
 QY 612 TTTATATTAGATATTTTACTTTTAAAGAGTCTTTAATAATTTT 654
 Db 69 TTTATATTAGATATTTTACTTTTAAATATCTTTAATAAAGTTT 27

RESULT 18
 AA039331/c
 LOCUS

AA039331 468 bp mRNA linear EST 10-MAY-1997


```

QY 265 AAGAGTTTACAGCTCTGAGCTGATTTGGGCTATGCTTACTGCTCCCTTTTGTCCTCA 324
D 393 AAGAGTTTACAGCTCTGAGCTGATTTGGGCTATGCTT----- 354
QY 325 GGAACCCACTGATGAAGAGCTTAAAGAGAGAAACACCCACTAAAGTTCGGAATCCCTAC 384
D 353 -GAACCCACTGATGAAGAGCTTAAAGAGAGAAACACCCACTAAAGTTCGGAATCCCTAC 295
QY 385 ACCTGATCAGAGAGAGAGATGATCAGGGTCAGCTGAGATTCAGTGCCTGACCTCGA 444
D 294 ACCTGATCAGAGAGAGAGATGATCAGGGTCAGCTGAGATTCAGTGCCTGACCTCGA 235
QY 445 AGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGGATGGATGTGAAGTGGTACTGA 504
D 234 AGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGGATGGATGTGAAGTGGTACTGA 175
QY 505 TGTCAGGGGAGAGATTCACCAAGAGAGAGACACTTTAAATGCCAGAGAGAGTGAAGG 564
D 174 TGTCAGGGGAGAGATTCACCAAGAGAGAGACACTTTAAATGCCAGAGAGAGTGAAGG 115
QY 565 GAAATCACAGCTTTAAAGGAGAGATGAGCTGAACACACAACTGTTTATATTAGATA 624
D 114 GAAATCACAGCTTTAAAGGAGAGATGAGCTGAACACACAACTGTTTATATTAGATA 55
QY 625 TTTTACTTTAAAGAGCTCTTAATAAATTTT 654
D 54 TTTTACTTTAAATATCTTAATAAAGTTT 25

RESULT 20
AA737311/c
LOCUS
DEFINITION
  n51904.s1 NCI CGAP Ew1 Homo sapiens cDNA clone IMAGE:1250166
  similar to TR:Q13066 Q13066 GAGE-2 PROTEIN. [1] ; mRNA sequence.
ACCESSION
  AA737311
VERSION
  1 (bases 1 to 532)
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 532)
  NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-re@mail.nih.gov
  Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D.,
  Ph.D.
  cDNA Library Preparation: David B. Krizman, Ph.D.
  cDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CCAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html
  Insert Length: 646 Std Error: 0.00
  Seq primer: -40m13 fwd. ET from Amersham
  High quality sequence stop: 429.
  Location/Qualifiers
    1..532
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:1250166"
      /tissue_type="Ewing's sarcoma"
      /lab_host="DH10B"
      /clone_lib="NCI CGAP Ew1"
      /note="Vector: pAMP10; mRNA made from Ewing's sarcoma,
      cDNA made by oligo-dT priming. Non-directionally cloned.
      Size-selected on agarose gel, average insert size 600 bp.
      Reference: Krizman et al. (1996) Cancer Research
      56:5380-5383."

```

```

BASE COUNT 127 a 128 c 107 g 169 t 1 others
ORIGIN
Query Match 56.4%; Score 377.6; DB 9; Length 532;
Best Local Similarity 52.1%; Fred. No. 2.7e-88;
Matches 417; Conservative 0; Mismatches 15; Indels 21; Gaps 1;
QY 215 TAGTGAATATGATTTGGCGAGGAAGATCAACATATAGGCCTAGGCCCAAGAAAGTTTA 274
D 432 TAGTGAATATGATTTGGCGAGGAAGATCAACATATAGGCCTAGGCCCAAGAAAGTTTA 373
QY 275 CAGCTCTCAGCTGATTTGGGCTATGCTTACTGGCTCCCTTTGTCGCCAGGACCCACT 334
D 372 CAGCTCTCAGCTGATTTGGGCTATGCTT-----GAACCCACT 334
QY 335 GATCAAGAGCCTTAAAGAGAGAAACACCCACTAAAGTCCGAATCCTACCTGATCAG 394
D 333 GATCAAGAGCCTTAAAGAGAGAAACACCCACTAAAGTCCGAATCCTACCTGATCAG 274
QY 395 AAGAGAGAGATGATCAGGTGCAGCTGAGATTCAGTGCCTGCTGAGGAGCGGATCTC 454
D 273 AAGAGAGAGATGATCAGGTGCAGCTGAGATTCAGTGCCTGCTGAGGAGCGGATCTC 214
QY 455 CAGGAGCTATGTCAGACAAAGACTGGGGATGGATGTGAAGTGGTACTGATGCAAGGG 514
D 213 CAGGAGCTATGTCAGACAAAGACTGGGGATGGATGTGAAGTGGTACTGATGCAAGGG 154
QY 515 AAGATTCACCAAGAGAGAGACTTTAAATGCCAGAGAGAGTGAAGGAAATCAGAG 574
D 153 AAGATTCACCAAGAGAGAGACTTTAAATGCCAGAGAGAGTGAAGGAAATCAGAG 94
QY 575 GTTTAAAGAGAGATGATGTAACACACAACTGTTTATATTAGATTTTACTTTTA 634
D 93 GTTTAAAGAGAGATGATGTAACACACAACTGTTTATATTAGATTTTACTTTTA 34
QY 635 AAGAGCTCTTAATAAATTTTGGCATGCTGCATC 667
D 33 CAATATCTCTTAATAAAGTTTAAAGCTTTTCTGTC 1

RESULT 21
H53211
LOCUS
DEFINITION
  YQ84C02.r1 Soares fetal liver spleen lNfLS Homo sapiens cDNA clone
  IMAGE:202466 5', mRNA sequence.
ACCESSION
  H53211
VERSION
  1 (bases 1 to 493)
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 493)
  Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman,
  M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
  Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisan, E., Waterston,
  R., Willamson, A., Wohlmann, F., and Wilson, R.
  The WashU-Merck EST Project
  Unpublished
  Contact: Wilson RK
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@watson.wustl.edu
  Insert Size: 663
  High quality sequence stops: 340
  Source: IMAGE Consortium, LLNL
  This clone is available royalty-free through LLNL; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  Insert Length: 663 Std Error: 0.00
  Seq primer: M13RPI
  High quality sequence stop: 340.

```


Db 83 AAGATAAGCTGAACACACACAACTGTTTTTATTATATAGATATTTTACTTTAAATAATCTTT 24

QY 644 AATAAATTTTT 654

Db 23 AATAAAGTTTT 13

RESULT 23

LOCUS W93043/c

DEFINITION ZD93f02.s1 Soares fetal_heart_NbHL19W Homo sapiens cDNA clone

ACCESSION W93043

VERSION W93043.1 GI:1422214

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1. (bases 1 to 512)

AUTHORS Hillier, L., Clark, N., Dubuque, T., Eliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

TITLE The WashU-Werck EST Project

JOURNAL Unpublished

COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 608 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 301.

FEATURES

Location/Qualifiers

1..512

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="GDB:1273595"

/db_xref="taxon:9606"

/clone="IMAGE:357051"

/sex="unknown"

/dev_stage="19 weeks"

/lab_host="DH10B (ampicillin resistant)"

/clone_lib="Soares fetal heart NbHL19W"

/note="Organ: heart; Vector: pT73D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCATCTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbHL19W."

BASE COUNT 119 a 121 c 108 g 162 t 2 others

ORIGIN

Query Match 55.6%; Score 372.2; DB 14; Length 512;

Best Local Similarity 90.8%; Pred. No. 6.9e-87;

Matches 471; Conservative 0; Mismatches 20; Indels 28; Gaps 6;

QY 143 CGCTCGGACTTTCTCT-CCTACTGAGAGC-CAGCCTATAGTCC--GCAGGCCAGTC 198

Db 511 CGGCTGGACTTTTCTCTCCCTACTGAGACGCCAGCCTATAGTCCCGCAGGCCAGTC 452

QY 199 CTCGCCAGGAACCTGAATAAGTGAATATGATGGTGGCAGGAAGA--TCAACATATAGGCT 256

Db 451 CTCCTCCGGAACGTAATTAGTGAATTATGATGGCCGAGGAAGATTCCACATATTGGCCT 392

QY 257 AGG-CCAGAGAAGTTTACAGCCTCCTGAGCTGATTTGGGCTATGCTTACTGGCTCCCC 315

Db 391 AGGCCCAAGAAGATTTACAGCCTCCTGAGCTGATTTGGGCTATGCTT-----343

QY 316 TTTGTCCAGGAACCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCCTAAAGTCG 375

Db 342 -----GAACCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCCTAAAGTCG 293

QY 376 GAATCCTACCTGATCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAGATGCC 435

Db 292 GAATCCTACCTGATCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAGATGCC 233

QY 436 TGACCTGGAAGCCGATCTCCAGAGCTATGTGAGACAAAGACTGGGGATGGATGTGAAGG 495

Db 232 TGACCTGGAAGCCGATCTCCAGAGCTATGTGAGACAAAGACTGGGGATGGATGTGAAGG 173

QY 496 TGGTACTGATGTCAGGGGAAGATTCTACCAAAAGCAGAGCATTAAATGCCGAGAGC 555

Db 172 TGGTACTGATGTCAGGGGAAGATTCTACCAAAAGCAGAGCATTAAATGCCGAGAGC 113

QY 556 AGGTGAAGGCAATCACAGCTTTAAAGGAAGATAAGCTCAAAACACACAAACTGTTTTTA 615

Db 112 AGGTGAAGGCAATCACAGCTTTAAAGGAAGATAAGCTCAAAACACACAAACTGTTTTTA 53

QY 616 TATTAGATATTTTACTTTAAAGAGTCTTAAATAATTTTT 654

Db 52 TATTAGATATTTTACTTTAAATAATCTTAAATAAGTTTT 14

RESULT 24

BE302172/c

LOCUS BE302172

DEFINITION BE302172.1 GI:9185920

similar to SW:GGE2_HUMAN Q13066 GAGE-2 PROTEIN. [1] ; mRNA sequence.

ACCESSION BE302172

VERSION BE302172.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1. (bases 1 to 445)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@p5b-romail.nih.gov
Tissue Procurement: ATCC
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
CDNA Sequencing by: Washington University Genome Sequencing Center
DNA distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml
Seq primer: -400P from Gibco
High quality sequence stop: 213.

FEATURES

Location/Qualifiers

1..445

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3048867"

/cell_line="MGC36"

/lab_host="DH10B"

/clone_lib="NIH MGC 10"

/note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 Kb. Library prepared by Life Technologies."

BASE COUNT 96 a 103 c 79 g 167 t

REFERENCE 1 (bases 1 to 437)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 570 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 407.
FEATURES
source
1..437
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2365214"
/lab_host="DH10B"
/clone_lib="Soares NSF F8 9W OT PA P S1"
/note="Organ: pooled; Vector: p773D-Pac (Pharmacia) with
a modified polylinker; Site: Not; Site_2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and clones: Soares NBHSP pool 1:
30384-310919, 323208-325895 Soares NB2HP pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares NB2HF8-9W pool 1:
758280-760583, 772104-774407 Soares NBHPA pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NBHOT
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares, and M. Fatima Bonaldo."
BASE COUNT 96 a 106 c 76 g 159 t
ORIGIN
Query Match 53.4%; Score 357.4; DB 9; Length 437;
Best Local Similarity 93.5%; Pred. No. 5.2e-83;
Matches 403; Conservative 0; Mismatches 6; Indels 22; Gaps 2;
Qy 224 ATGAGTTGGGAGAGATCAACATATAGGCTAGGCGCAGAGAGAGTTTACAGCTCTCT 283
Db 437 ATGAGTTGGGAGAGATC-ACCTATAGGCTAGGCGCAGAGAGAGTTTACAGCTCTCT 379
Qy 284 GAGCTGATTGGGGCTATGCTTACTGCTCCCTTGTCCAGGAAACCCACTGATGAAGAG 343
Db 378 GAGCTGATTGGGGCTATGCTT-----GAAACCCACTGATGAAGAG 340
Qy 344 CCTAAGAGAGAGAACACCCACTAAAGTCGGAATCTTACCTGATCAGAGAGAGAA 403
Db 339 CCTAAG 280
Qy 404 GATGATCAGGTCGAGCTGAGATTCAAGTGCCTCAGCTGAGGAGCGATCTCCAGGAGCTA 463
Db 279 GATGATCAGGTCGAGCTGAGATTCAAGTGCCTCAGCTGAGGAGCGATCTCCAGGAGCTA 220
Qy 464 TGTGAGCAAAAGACTGGGGATGGAATGGAAGGTGGTACTGATGTCAAGGGGAGAGTTCTTA 523
Db 219 TGTGAGCAAAAGACTGGGGATGGAATGGAAGGTGGTACTGATGTCAAGGGGAGAGTTCTTA 160
Qy 524 CCAAAGCAGAGAGCACTTTTAAATGCGAAGCAGGTGAAGGGAATCACAGGTTTAAAGG 583
Db 159 CCAAAGCAGAGAGCACTTTTAAATGCCAGAGCAGGTGAAGGGAATCACAGGTTTAAAGG 100
Qy 584 AAGTAAGCTGAACACACAACTGTTTTTATATTAGATATTTTACTTTTAAAGAGTCCTT 643
Db 99 AAGTAAGCTGAACACACAACTGTTTTTATATTAGATATTTTACTTTTAAATATCTTT 40
Qy 644 AATAAATTTT 654
|||||

Db 39 AATAAAGTTTT 29
RESULT 27
T87104
LOCUS
DEFINITION
T87104
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
T87104 555 bp mRNA linear EST 17-MAR-1995
IMAGE:115340 5', mRNA sequence.
T87104 GI:715456
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman
M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston
R., Williamson, A., Wohlmann, P. and Wilson, R.
The WashU-Merck EST Project
Unpublished
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 400
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: M13RP1
High quality sequence stop: 400.
FEATURES
source
1..555
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:470957"
/db_xref="taxon:9606"
/clone="IMAGE:115340"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen LNLs"
/note="Organ: liver and spleen; Vector: p773D (Pharmacia)
with a modified polylinker; Site: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACTGGAAGATTAAATAAGATCTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 147 a 126 c 152 g 127 t
ORIGIN
Query Match 52.6%; Score 351.8; DB 14; Length 555;
Best Local Similarity 87.6%; Pred. No. 1.6e-81;
Matches 496; Conservative 0; Mismatches 35; Indels 35; Gaps 9;
Qy 14 CCGACTTCAGTGTGATGTTCTTGGACACCTGCTGCTGATGTTCTCTGGGATC 73
Db 5 CAGCTTCAGTGTGATGTTCTTGGACACCTGCTGCTGATGTTCTCTGGGATC 64
Qy 74 TTCCCTTCGACCCCTTTCGCCACGTTGGAGCTGGAGGTGTGAGGGGC 133
Db 65 TTCCCTTCGACCCCTTTCGCCACGTTGGAGCTGGAGGTGTGAGGGGC 124
Qy 134 AGTTTCCAGCGCTGAGACTCTTTCTCTACTGAGACGAGCCTATAGTCCGAGC 193
Db 125 AGTTTCCAGCGCTGAGACTCTTTCTCTACTGAGACGAGCCTATAGTCCGAGC- 183

	RESULT	32	
REFERENCE	W93171		
AUTHORS	Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston R., Williamson,A., Wohldmann,P. and Wilson,R.		
TITLE	The WashU-Merck EST Project		
JOURNAL	Unpublished		
COMMENT	Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.wustl.edu This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert Length: 608 Std Error: 0.00 Seq primer: mob.REGA+ET High quality sequence stop: 336. Location/Qualifiers		
FEATURES	source		
	1..401		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="GBB:3771597"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:202466"		
	/sex="male"		
	/dev_stage="20 week-post conception fetus"		
	/lab_host="DH10B (ampicillin resistant)"		
	/clone_lib="Soares fetal liver spleen INFLS"		
	/notes="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site 1: Pac I; Site 2: Eco RI; 1st strand cDNA was primed with a Pac I oligo(dT) primer [5' - AACTGGAAGATTAATAAGAATCTTTTTTTTTTTTTT 3'] , double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Patima Bonaldo."		
	BASE COUNT	85 a	98 c 72 g 141 t 5 others
	ORIGIN		
	Query Match	47.4%; Score 316.8;	DB 14; Length 401;
	Best Local Similarity	97.6%;	Pred. No. 2.4e-72;
	Matches	321; Conservative	0; Mismatches 8; Indels 0; Gaps 0;
QY	326	GAAACCCACTGTAGGAGCGCCTTAAAGAGAGAACACCACCACTAAAAAGTCGGATCCTACA	385
Db	336	GAAACCCACTGTAGGAGCGCCTTAAAGAGAGAACACCACCACTAAAAGTCGGATCCTACA	277
QY	386	CCTGATCAGAGAGAGAAAGTAGATCATGGGTGCAGCTGAGATTCAAAGTGCTGCACCTGGAA	445
Db	276	CCTGATCAGAGAGAGAAAGTAGATCATGGGTGCAGCTGAGATTCAAAGTGCTGCACCTGGAA	217
QY	446	GCCGNTCTCCAGGAGCTATGTCACAGAAAGACTGGGATGGATGTGAAGTGCTGATCTGAT	505
Db	216	GCCGATCTCCAGGAGCTATGTCACAGAAAGACTGGGATGGATGTGAAGTGCTGATCTGAT	157
QY	506	GTCAGGGGGAAGATTCTTACCAAAAGCGAGCACATTTTAAATGCCAGAGCAGGTGAAGGG	565
Db	156	GTCAGGGGGAAGATTCTTACCAAAAGCGAGCACATTTTAAATGCCAGAGCAGGTGAAGGG	97
QY	566	AAATCACAGGTTTAAAGGAAGATAGCTGAAACAAACAACACTGTCTTTATATTAGATAT	625
Db	96	AAATCACAGGTTTAAAGGAAGATAGCTGAAACAAACAACACTGNTTTTATATTAGTAT	37
QY	626	TTTACTTTAAAGAGCTTTAAATAAATTTT	654
Db	36	TTTAAAGTAAATACTTTAAATAAAGTTT	8

```

182 CGAGGAGATCATATAGCTTAGG-CAAGAGAGATTACAGCTCTCTGAGTGAT 240
183 |||||
292 TGGGCTATGCTTACTGCTCCCTTTGTCCTCCAGAACCCACTGATGAAGAGCCTTAAGA 351
293 |||||
241 TGGGCTATGCTT-----GAACCCACTGATGAAGAGCCTTAAGA 279
242 |||||
352 AGAGAAACCCACCCTAAAGTCGGAATCCCTACACCTGATCAGAAAGAGAGATGATCA 411
353 |||||
280 AGAGAAACCCACCCTAAAGTCGGAATCCCTACACCTGATCAGAAAGAGAGATGATCA 339
281 |||||
412 GGGTCAGCTG-AGATTCAAGTCGCTGACCT-GGAAGCCGATCTCCAGGA 459
413 |||||
340 GGGTCAGCTGAAGATTCAAGTCGCTGACCTGGGAAGCCGATCTCCAGGA 389
341 |||||

RESULT 33
N42010 420 bp mRNA linear EST 24-JAN-1996
LOCUS YW69f02.r1 Soares placenta 8to9weeks 2NBHP8to9W Homo sapiens CDNA
DEFINITION clone IMAGE:257499 5', mRNA sequence.
ACCESSION N42010
VERSION N42010.1 GI:1166041
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 420)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
R., Williamson,A., Wohlmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 321
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Putative full length read
Seq primer: T7
High quality sequence stop: 321.
Location/Qualifiers
1. .420
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3887109"
/db_xref="taxon:9606"
/clone="IMAGE:257499"
/dev_stage="two placentae: one from 8 weeks and another
from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares placenta 8to9weeks 2NBHP8to9W"
/notes="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(37) primer [5'
TGTTACCAATCTAAGTGGAGCGCGCGGATTTTTTTTTTTTTTTT 3'].
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo."

BASE COUNT 121 a 91 c 123 g 81 t 4 others
ORIGIN
Query Match 46.8%; Score 313.4; DB 14; Length 420;
Best Local Similarity 87.3%; Pred. No. 1.9e-71;

182 CGAGGAGATCATATAGCTTAGG-CAAGAGAGATTACAGCTCTCTGAGTGAT 240
183 |||||
292 TGGGCTATGCTTACTGCTCCCTTTGTCCTCCAGAACCCACTGATGAAGAGCCTTAAGA 351
293 |||||
241 TGGGCTATGCTT-----GAACCCACTGATGAAGAGCCTTAAGA 279
242 |||||
352 AGAGAAACCCACCCTAAAGTCGGAATCCCTACACCTGATCAGAAAGAGAGATGATCA 411
353 |||||
280 AGAGAAACCCACCCTAAAGTCGGAATCCCTACACCTGATCAGAAAGAGAGATGATCA 339
281 |||||
412 GGGTCAGCTG-AGATTCAAGTCGCTGACCT-GGAAGCCGATCTCCAGGA 459
413 |||||
340 GGGTCAGCTGAAGATTCAAGTCGCTGACCTGGGAAGCCGATCTCCAGGA 389
341 |||||

Matches 385; Conservative 0; Mismatches 30; Indels 26; Gaps 3;
QY 101 TCACCGCTGGGAGCTGTGAGAGTGTGAGGGGACAGTTCCAGCCGCTCTGGAGCTTTTCTC 160
Db 1 TCACCGCTGGGAGCTGTGAGAGTGTGAGGGGACAGTTCCAGCCGCTCTGGAGCTTTTCTC 60
QY 161 TCTACTGAGACGCGACCTTATAGTTCGCGAGCCAGTCTCTCCAGGAACTGAAATAGTGA 220
Db 61 TCTACTGAGACGCGACCTTATAGTTCGCGAGCCAGTCTCTCCAGGAACTGAAATAGTGA 120
QY 221 AATATGAGTTGGCGAGGAGATCAACATATAGGCTAGGCCCAAGAGAGTTTACAGCCT 280
Db 121 AATATGAGTTGGCGAGGAGATCAACATATAGGCTAGGCCCAAGAGAGTTTACAGCCT 180
QY 281 CTTGAGCTGATTGGGGCTATGCTTACTGCTCCCTTTGTTCCCGAGAACCCCTACTGATGA 340
Db 181 CTTGAGCTGATTGGGGCTATGCTT-----GAACCCACTGATGA 219
QY 341 GAGCCTAAAGAGAGAGAACCCCTAAAGTCGGAATCCCTACACCTGATCAGAAAGAGA 400
Db 220 GAGCCTAAAGAGAGAGAACCCCTAAAGTCGGAATCCCTACACCTGATCAGAAAGAGA 279
QY 401 GAAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTCGAAAGCCGATCTCCAGGAG 460
Db 280 GAAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTCGAAAGCCGATCTCCAGGAG 339
QY 461 CTATG-----TCAGACAAAGACTGGGATCGATGTGAAGT-GGTACTGATGTCAAGGGGA 515
Db 340 CTATGTCAGACCAAGACTGGGGGATGATGTGAAGTGGGTACTGATGTCAAGGGGG 399
QY 516 AGATTCTACCAAAAGCAGAGC 536
Db 400 AAGNATTCCTANCAAAAAGC 420

RESULT 34
N40147 547 bp mRNA linear EST 22-JAN-1996
LOCUS YW73e12.r1 Soares placenta 8to9weeks 2NBHP8to9W Homo sapiens CDNA
DEFINITION Clone IMAGE:257902 5', mRNA sequence.
ACCESSION N40147
VERSION N40147.1 GI:1163692
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 547)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
R., Williamson,A., Wohlmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 450
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: T7
High quality sequence stop: 450.
Location/Qualifiers
1. .547
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3887512"
/db_xref="taxon:9606"
/clone="IMAGE:257902"

FEATURES
source

```

```

/dev_stage="two placentae: one from 8 weeks and another
from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="8oaves.placenta.8to9weeks.2NbHP8to9W"
/note="Organ: placenta; Vector: p77t3D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTACATCTGATGGGAGCGCGCGATTTTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p77t3 vector
(Pharmacia). Library constructed by Bento Soares and
M.Patina Bonaldo."
159 a 111 c 147 q 123 t 7 others
BASE COUNT

```

	BASE COUNT	159 a	111 c	147 g	123 t	7 others
	ORIGIN					
	Query Match	45.6%	Score 304.8;	DB 14;	Length 547;	
	Best Local Similarity	75.7%;	Pred. No. 3.8e-69;			
	Matches	Conservative 0;	Mismatches 83;	Indels 58;	Gaps 3;	
Qy	77	CCTTCGACCCCTTTGCCACAGTGTGACCGCTGGGGAGCTGTGAGAGTGTCAGGGGCACG	136			
Dd	1	CCCTNGCNATNTTCCACAGTGTGAATGCCCCGTGGAGTGTGAGGTTGTGAGGTCGCGGTCGG	60			
Qy	137	TTCACGCGCTCGACTCTTCTCTCTPACTAGACGCAGCCTATAGTCCGCAAGCCAG	196			
Dd	61	TTCCTGCTGTCTGGACTTTTTCTGTCCCACCTGACAGCCAGCTGT-----	104			
Qy	197	TCCTCCCAGGAACTCGAATAGTAGTAATATGATTGTGGCGAGAGAAGTCAACATATAGGCCT	256			
Dd	105	-----GTGAANTATGATTTGGCGAGGAGATCAACATATAGGCCT	144			
Qy	257	AGGCCAAGAAGTTTACAGCCTCCTGAGCTGTATTGGGGCTATGCTTACTGGCTCCCTCT	316			
Dd	145	AGGCGAAGAGAAGTGTACCACCTCCTGAGCTGTATTGGGCCCTATGCT-----	191			
Qy	317	TGTGCCCAAGAACCCACCTGTATGAACAGCCTAAAGAGAGAACAACCCACCTAAAAAGTCGG	376			
Dd	192	-----GGACCCCGGTGATGAGAGGCCTCAGCAAGAGGAACCACTGAAAGTCGG	243			
Qy	377	AATCTTACACCTGTATCAGAAGAGAGAAGATGATCAGGTGCGAGCTGAGATTCAGATGCCT	436			
Dd	244	GATCTGCACTTGGTCAGSAGAGAAGAAGATCAGGGTGCAGCTGAGACTCAAGTGCCT	303			
Qy	437	GACCTGGAAGCCGATCTCCAGAGAGCTATGTCAGACAAGAGACTGGGATGGATGTCAAAGT	496			
Dd	304	GACCTGGAAGCTGATCTCCAGAGAGCTGCTCASTCAAGACTGGGGTGAATGTGMAAT	363			
Qy	497	GGTACTGTGTTCAAGGGGAAGATTCTACAAAGCGAGACATTTTAAATGCCAGAGCA	556			
Dd	364	GGTCTGTGATCACAGGGGAAGATTCTGCCAAATCAGAACAAATTTAAAATGCCAGAGCA	423			
Qy	557	GCTGAGGGAATACACAGGTTTTAAAGGAAGATAAGCTGAAAACAAC-AACTGTTTTTA	615			
Dd	424	GGTGACAGCAACCAAGTTTTAATGAAGACAAGCTGAAAACAACAAAACGTGTTTTA	483			
Qy	616	TATTAGATATTTACTTTAAAGAGCTCTTAATAATTTTTTG	655			
Dd	484	TCTAGATATTTGACTTTNAAAATATCGANATTAACCTTTTG	523			

RESULT 35
R99925/c
LOCUS
DEFINITION
R99925 410 bp mRNA linear EST 14-SEP-1995
YQ75602.s1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:201723 3', mRNA sequence.
ACCESSION
R99925
VERSION
R99925.1 GI:986526
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Chordata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa;
Mammalia; Eutheria; Primates;
Carnivora; Canidae; Canis; Canis lupus familiaris; Homo.

REFERENCE
AUTHORS

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 1, pp. 1-10
2. The Impact of Technology on Education	Journal of Educational Research	1980, Vol. 83, No. 2, pp. 11-20
3. The Importance of Parental Involvement	Journal of Educational Research	1980, Vol. 83, No. 3, pp. 21-30
4. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 4, pp. 31-40
5. The Role of the School in the Community	Journal of Educational Research	1980, Vol. 83, No. 5, pp. 41-50
6. The Impact of Teacher Education on Classroom Practice	Journal of Educational Research	1980, Vol. 83, No. 6, pp. 51-60
7. The Importance of Teacher Collaboration	Journal of Educational Research	1980, Vol. 83, No. 7, pp. 61-70
8. The Effect of Teacher Attitudes on Student Behavior	Journal of Educational Research	1980, Vol. 83, No. 8, pp. 71-80
9. The Role of the School in the Development of the Child	Journal of Educational Research	1980, Vol. 83, No. 9, pp. 81-90
10. The Impact of Teacher Training on Student Learning	Journal of Educational Research	1980, Vol. 83, No. 10, pp. 91-100

1 (bases 1 to 410)

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, B., Rohlfing, T., Soares, R., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohldmann, P., and Wilson, R.

The WashU-Merck EST Project

Unpublished

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Insert Size: 1596

High quality sequence stops: 361

Source: IMAGE Consortium, LNL

This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 1596 Std Error: 0.00

Seq primer: Promega -21mi3

High quality sequence stop: 361.

Location/Qualifiers

1. 410

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="GDB:3770854"

/db_xref="taxon:9606"

/clone="IMAGE:201723"

/sex="male"

/dev_stage="20 week-post conception fetus"

/lab_host="DH10B (ampicillin resistant)"

/clone_lib="Soares fetal liver spleen cDNAs"

/note="Organ: Liver and Spleen; Vector: pVT73D (Pharmacia) with a modified polylinker; Site 1: Pac I; Site 2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5', AACTGGAGCAATTAATAAGACTTTTTTTTTTTTTTTT 3', double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pVT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

91 a 98 c 75 g 140 t 6 others

BASE COUNT	91 a	98 c	75 g	140 t	6 others
ORIGIN					
Query Match	45.34;	Score	303.2;	DB 14;	Length 410;
Best Local Similarity	89.7%;	Pred. No.	9.1e-69;		
Matches 376;	Conservative	0;	Mismatches 18;	Indels 25;	Gaps 4;
Qy	240	GATCAACATATAGCGCTAGGCCCAAGAAAGTTTACAGCCTCCTGA--GCTGATTGGGCG	297		
Db	410	GATCAACATATAGCGCTAGGCCCAAGAAAGTTTACAGCCTCCTGAGNCTGATTGGGCG	351		
Qy	298	TATGTTTACTGGCTCCCTTTGTGCCAGGAACCCACTGATGAAGAG--CCTTAAGAGAGA	356		
Db	350	TATGCTT-----GAAACCCTGNTGAAGAGCCCTTAAGNAGAGA	312		
Qy	357	AACCAACCCACTAAAGTTCGGAATCCTACACCTGATCAGAAGAGAGAAGATGATCAGGGTG	416		
Db	311	AACCAACCCACTAAAGTTCGGAATCCTACACCTGATCAGAAGAGAGAAGATGATCAGGGTG	252		
Qy	417	CAGCTGAGATTCAGTGCCTGA--CCTGGAGCCGATCTCCAGGAGCTATGTCAGACAAG	475		
Db	251	CAGCTGAGATTCAGTGCCTGANCCTGAGGCCGATCTCCAGGAGCTATGTCAGACAAG	192		
Qy	476	ACTGGGGATGGATGTGAAGGTGGTACTGATGTCAAGGGGAAGATTCCTACAAAGCAGAG	535		
Db	191	ACTGGGGATGGATGTGAAGGTGGTACTGATGTCAAGGGGAAGATTCCTACAAAGCAGAG	132		
Qy	536	CACTTTAAATGCCAGAGCAGGTGAAGGGAAATCAACAGGTTTTAAAGCAAGATAAGCTGA	595		
Db	131	CNCTTTAAATGCCAAGAGCAGGTGAAGGGAAATCAACAGGTTTAAAGCAAGATAAGCTGA	72		
Qy	596	AACAACACAAACTGTTTTTATATATAGATPATTTTACTTTTAAAGAGTCTTATAAATTTTT	654		

```

71 AACACACAAACTGTTTATATTAGATATTTTAAAGTAAATCTTATTAAGTTTT 13
|||||
Db 71 AACACACAAACTGTTTATATTAGATATTTTAAAGTAAATCTTATTAAGTTTT 13
|||||

RESULT 36
CB992682 570 bp mRNA linear EST 01-MAY-2003
LOCUS AGENCOURT 13633348 NIH_MGC_148 Homo sapiens cDNA clone
DEFINITION IMAGE:30338928 5', mRNA sequence.
ACCESSION CB992682
VERSION CB992682.1 GI:30287202
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 570)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM366 row: g column: 01
High quality sequence stop: 558.
Location/Qualifiers
1. 570
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30338928"
/tissue_type="pre-eclamptic placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_148"
/notes="Organ: Placenta; Vector: pBluescriptR; Site 1:
all-XhoI; Site 2: BamH; Library is oligo-dT primed and
directionally cloned using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert
size 2.3 kb and normalized to ROT 5. This is a primary
library enriched for full-length clones and constructed
using the Cap-trapper method (Carninci, in preparation).
Library constructed by M. Brownstein (NIMH/NHGRI).
National Institutes of Health." Note: this is a NIH_MGC
Library."
BASE COUNT 174 a 116 c 157 g 123 t
ORIGIN
Query Match 44.3%; Score 295.4; DB 14; Length 570;
Best Local Similarity 76.2%; Pred. No. 6.2e-67;
Matches 430; Conservative 0; Mismatches 76; Indels 58; Gaps 3;

QY 93 CCACGTGTGACCGCTGGGAGCTGTGAGAGTGTGAGGGGACGCTTCAGCGCTGTGAC 152
Db 24 CCAGGCTGTGAATGCCCTGGAGTGTGAGGGTGTGAGGGTCCGCTTCGCTGTGAC 83
QY 153 TCTTTCTCTCTACTGAGCGAGCGAGCTATAGTCCGAGCGGCGCTCCCGAGGACTGA 212
Db 84 TTTTCTGTCCCACTGAGCGAGCTGT----- 111
QY 213 AATAGTGAATATGAGTTGGCGAGGAAGATCAACATATAGGCTAGGCCAAGAAAGTT 272
Db 112 ----GTGAATATGATTTGGCGAGGAAGATCAACATATAGGCTAGGCCGAGGAGAGTG 167
QY 273 TACAGCTCTGAGCTGATGGGCTAGCTTACTGGCTCCCTTTGTCACGAGAACCA 332
Db 168 TACCACCTCCTGAGCTGATGGGCTATGCT-----GGAGCGCG 206

```

```

QY 333 CTGATGAAGAGCCTAAAGAGAGAGAAACCCACCTAAAGTCGGAATCCTACACCTGATC 392
Db 207 GTGATGAGAGCCTCAGCAAGAGAGAAACCCACCACTGAAAGTCGGGATCTCTCACCTGTC 266
QY 393 AGAAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAGTGCCTGACCTGGAAGCCGATC 452
Db 267 AGGAGAGAGAGAGATCAGGGTGCAGCTGAGATTCAGTGCCTGACCTGGAAGCTGATC 326
QY 453 TCCAGGAGCTATGTCAGACAAAGACTGGGGATGGATGTGAAAGTGGTACTGATGTCAGG 512
Db 327 TCCAGGAGCTGTCTCAGTCAAAGACTGGGGTGAATGTGGAATGGTCTCTGATGACCAAG 386
QY 513 GGAAGATTTCTACCAAGACGAGACCTTTAAATGCCAGAGCAGGTGGAAGGAAATCAC 572
Db 387 GGAAGATTTCTGCCAAATCAATCAAAATTTAAATGCCAGAGGAGGTGACAGGAACAC 446
QY 573 AGGTTTAAAGAGAGATAAGCTGAAACACAC-AAACTGTTTTTATATTAGATATTTTACT 631
Db 447 AGGTTTAAATGAGACAGCTGAAACACACAACTGTTTTTATTAAGATATTGACT 506
QY 632 TTAAGAGCTCTTAATAAATTTTG 655
Db 507 TAAATATCAATAAATTTTG 530

CB994958 570 bp mRNA linear EST 01-MAY-2003
AGENCOURT 13643389 NIH_MGC_148 Homo sapiens cDNA clone
IMAGE:30335086 5', mRNA sequence.
ACCESSION CB994958
VERSION CB994958.1 GI:30289478
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 570)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM356 row: f column: 23
High quality sequence stop: 570.
Location/Qualifiers
1. 570
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30335086"
/tissue_type="pre-eclamptic placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_148"
/notes="Organ: Placenta; Vector: pBluescriptR; Site 1:
all-XhoI; Site 2: BamH; Library is oligo-dT primed and
directionally cloned using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert
size 2.3 kb and normalized to ROT 5. This is a primary
library enriched for full-length clones and constructed
using the Cap-trapper method (Carninci, in preparation).
Library constructed by M. Brownstein (NIMH/NHGRI).
National Institutes of Health." Note: this is a NIH_MGC
Library."
FEATURES
source
REFERENCE 1 (bases 1 to 570)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM356 row: f column: 23
High quality sequence stop: 570.

```

```
BASE COUNT      175 a      115 c      157 g      123 t
ORIGIN
Query Match      44.3%; Score 296.4; DB 14; Length 570;
Best Local Similarity 76.2%; Pred. No. 6.2e-67;
Matches 430; Conservative 0; Mismatches 76; Indels 58; Gaps 3;
QY 93 CCAGGTGTCACCGCTGGGAGCTGTGAGAGTGTGAGGGGCAAGTTCACAGCGCTCTGGAC 152
Db 24 CCAGGGTGTGAATGCCCTGAGTTGTGAGGTGTGAGGGTTCGCTCTCTGGAC 83
QY 153 TCCTTCTCTCTACTAGAGCGCAGCTATAGTCCGCGAGCCAGTCTCTCCAGAACTGA 212
Db 84 TTTTCTGTCCCACTGAGACGAGCTGT----- 111
QY 213 AATAGTAAATATAGTTGGCGAGGAAGATCAACATATAGGCTAGGCCAAGAAAGTT 272
Db 112 ----GTGAATATAGTTGGCGAGGAAGATCAACATATAGGCTAGGCCGAGGAAAGTG 167
QY 273 TACAGCCTCTGAGCTGATTTGGGGCTATGCTTACTGCTCCCTTTGTCCAGGAACCCA 332
Db 168 TACACCTCTGAGCTGATTTGGGCTATGCT-----GGAGCCCG 206
QY 333 CTGATGAAGAGCTTAAAGAGAGAAACACCGACTAAAGTCGAATCCTACACCTGATC 392
Db 207 GTGATGAGGAGCTTCAGCAAGAGAAACCACTGAAGTCGGGATCCTGCACCTGGTC 266
QY 393 AGAAGAGAGAAAGTATGATCAGGTGTCAGCTGAGATTCAGTGCCTGAGCGCGATC 452
Db 267 AGNAGAGAGAAAGATCAGGGTGCAGCTGAGACTCANGTCCCTGACCTGGAGCTGATC 326
QY 453 TCCAGGAGCTATGTCAGCAAAAGACTGGGATGATGTGAAGTGTGATCTGATGTCAGG 512
Db 327 TCCAGGAGCTGTCTCAGTCAAGACTGGGGTGAATGTGAAATGTCTCTGATGACCA 386
QY 513 GGAAGATTTCTACCAAGAGAGAGACACTTAAATGCCAGAGAGAGGTGAAGGAATCAC 572
Db 387 GGAAGATTTCTCCAAATTCAGAACATTTAAATGCCAGAGAGAGGTGACAGCAACCA 446
QY 573 AGGTTTAAAGAGAGATAGCTGAACCAACAC- AAACCTGTTTTATATAGATATTTACT 631
Db 447 AGGTTTAAATGAGACAGCTGAACCAACCAACAACTGTTTTATCTAAGATATTTGACT 506
QY 632 TTAAGAGCTCTTAAATATTTTG 655
Db 507 TAAATATCAAAATAAACCTTTG 530

RESULT 38
BG354572      499 bp      mRNA      linear      EST 05-MAR-2001
LOCUS      BG354572
DEFINITION      PLAC6 Human placenta cDNA expression library Homo sapiens CDNA
clone 253708 similar to PLAC6 Placenta specific gene 6, mRNA
sequence.
ACCESSION      BG354572
VERSION      BG354572.1 GI:13198770
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      1 (bases 1 to 499)
AUTHORS      Wong, S.L. and Walker, M.G.
TITLE      Placental genes identified by large-scale expression analysis
JOURNAL      Unpublished
COMMENT      Contact: Michael Walker
Incyte Genomics
3160 Porter Drive, Palo Alto, CA 94304, USA
Tel: 650 845-5771
Fax: 650 621 8514
Email: mwalker@incyte.com.
Location/Qualifiers
1. .499

FEATURES
Source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="253708"
/clone_lib="Human placenta cDNA expression library"
/notes="Organ: placenta"
BASE COUNT      148 a      100 c      139 g      112 t
ORIGIN
Query Match      43.4%; Score 290.2; DB 10; Length 499;
Best Local Similarity 76.9%; Pred. No. 2.5e-65;
Matches 419; Conservative 0; Mismatches 68; Indels 58; Gaps 3;
QY 112 GAGCTGTGAGAGTGTGAGGGGCAAGTTCACAGCGCTCTGGACTCTCTCTCTACTGAGA 171
Db 1 GAGTTGTGAGGGTGTGAGGGTCTGCTCTCTCTGGACTCTCTCTCTCTCTCTCTCTGAGA 60
QY 172 CGAGCCTATAGGTCCGAGGCCAGTCTCTCCAGAACTGAAATAGTGAATATGAGTTG 231
Db 61 CGAGCTGT-----GTGAATATGATTTG 84
QY 232 CGAGGAAGATCAACATATAGGCTAGGCCAAGAAAGTTTACAGCCTCTCTGAGCTGAT 291
Db 85 CGAGGAAGATCAACATATAGGCTAGGCCGAGGAAGTGTACACCTCTCTGAGCTGAT 144
QY 292 TGGGCTATGCTTACTGCTCCCTTTGTCCAGGAACCCCACTGATGAAGAGCCTAAAGA 351
Db 145 TGGGCTATGCT-----GGAGCCCGTGTGAGGAGCCTCAGCA 183
QY 352 AGAGAAACACCCACTAAAGTCGAAATCCTACACCTGATCAGAGAGAGAGATGATCA 411
Db 184 AGAGAAACACCAACTGAAAGTCGGGATCCTGACCTGTCAGGAGAGAGAGATGATCA 243
QY 412 GGGTCAGCTGAGATTCAAAGTGCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTCAG 471
Db 244 GGGTCAGCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCAGTC 303
QY 472 AAAGACTGGGATGATGTGAAGTGTGATGTCTCAAGGGGAGATTTCTACCAAAAGC 531
Db 304 AAAGACTGGGATGATGTGAAGTGTGATGTCTCAAGGGGAGATTTCTGCCAAATC 363
QY 532 AGAGCACTTTAAATGCCAGAGAGAGGTGAAGGGAAATCACAGGTTTAAAGAGAGATAAG 591
Db 364 AGAACAAATTTAAATGCCAGAGAGAGGTGACAGGCAACACAGGTTTAAATGAAGACA 423
QY 592 CTGAAACAAACAC- AAACCTGTTTTATATAGATATTTTACTTTAAAGAGCTCTTAATAAT 650
Db 424 CTGAAACAAACAAACCTGTTTTATCTAAGATATTTGACTTAAATAATCAAAATAAAC 483
QY 651 TTTTG 655
Db 484 TTTTG 488

RESULT 39
CB961487      520 bp      mRNA      linear      EST 29-APR-2003
LOCUS      CB961487
DEFINITION      AGENCOURT 13894159 NIH.MGC.148 Homo sapiens cDNA clone
IMAGE:30348215 5', mRNA sequence.
ACCESSION      CB961487
VERSION      CB961487.1 GI:30217604
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      1 (bases 1 to 520)
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE      NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
```

cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM390 row: 4 column: 24
High quality sequence stop: 520.
Location/Qualifiers
1. 524

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30348215"
/tissue_type="pre-eclamptic placenta"
/lab_host="DH10B TONa"
/clone_lib="NIH_MGC_148"
/note="Organ: placenta; Vector: pBluescriptR; Site_1: all-XhoI; Site_2: BamHI; Library is oligo-dT primed and directionally cloned using primer 5'-TTTTTATTTTATTTT-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 174 a 105 c 134 g 107 t

Query Match 42.8%; Score 286.4; DB 14; Length 520;
Best Local Similarity 82.3%; Pred. No. 2.6e-64;
Matches 362; Conservative 0; Mismatches 56; Indels 22; Gaps 2;

QY 217 GTGAATATGAGTGGCGAGGAAGATCAACATATAGGCTAGGCCAAGAGAGAGTTTACA 276
DB 58 GTGAATATGAGTGGCGAGGAAGATCAACATATAGGCTAGGCCAAGAGAGTTTACC 117
QY 277 GCTCTGAGCTGATGGGGCTATGCTTACTGCTGCCCTTGTCTCCAGGAACCCACTGA 336
DB 118 ACCTCTGAGCTGATGGGCTATGCT-----GGAGCCCGGTGA 156
QY 337 TGAAGAGCTTAAGAGAGAAACCCACCTAAAGTCGGAATCCTACACCTGATCAGAA 396
DB 157 TGAGGAGCTCAGCAGAGAGAACCCACCTAAAGTCGGAATCCTGACCTGGTCAGGA 216
QY 397 GAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAGCGGATCTCCA 456
DB 217 GAGAGAGAGATCAGGGTGCAGCTGAGACTCAAGTGCCTGACCTGGAGCTGATCTCCA 276
QY 457 GGAGCTATGTCAGACAAAGACTGGGGTGGATGGATGGATGGATGGATGGATGGATGGAT 516
DB 277 GGAGCTGTCTCATCAAGACTGGGGTGGATGGATGGATGGATGGATGGATGGATGGAT 336
QY 517 GATTCTACCAAGCAGAGACCTTTTAAATGCCAAGCAGGTGAGGAGGAATCAGGT 576
DB 337 GATTCTGCCAAATCAGAACATTTTAAATGCCAAGGAGGTGACAGGCAACCCAGGT 396
QY 577 TTAAGGAGATGAGCTGAACACAC-AAAATGTTTTTATATTAGATTTTACTTTAA 635
DB 397 TTAATGAAGACAGCTGAACACACAAACAACTGTTTTTATTAAGATATTGACTTAA 456
QY 636 AGAGTCTTTAATAATTTTG 655
DB 457 AATATCAAAATAAATTTTG 476

RESULT 40
AF318372
LOCUS
DEFINITION Homo sapiens pp9012 mRNA, complete cds.
ACCESSION AF318372

AF318372.1 GI:18027835

HTC.
Homo sapiens (human)
Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 524)
Zhou,X.M., Zhang,P.P., Jiang,H.Q., Huang,Y., Qin,W.X., Zhao,X.T., Wan,D.F. and Gu,J.R.
Novel human cDNA clones with function of inhibiting cancer cell growth
Unpublished
2 (bases 1 to 524)
Zhou,X.M., Zhang,P.P., Jiang,H.Q., Huang,Y., Qin,W.X., Zhao,X.T., Wan,D.F. and Gu,J.R.
Direct Submission
Submitted (02-NOV-2000) National Laboratory For Oncogenes & Related Genes, Shanghai Cancer Institute, 25/In 2200 Xie-Tu Road, Shanghai 200032, P. R. China
Location/Qualifiers
1. 524
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. 524
/gene="pp9012"
73. 408
/gene="pp9012"
/codon_start=1
/product="unknown"
/protein_id="AAL55879.1"
/db_xref="GI:18027835"
/translation="MWRGRSTYRHRPRRSVPPBLIGFMLEFGDEEQBPPTESR DPAPQEREDQAAETQVPLEADLQLBSQSGTCGCGNPDQKILPKSEQPMP EGDQRPQV"
BASE COUNT 177 a 99 c 138 g 110 t

Query Match 42.6%; Score 285.2; DB 11; Length 524;
Best Local Similarity 82.0%; Pred. No. 5.3e-64;
Matches 361; Conservative 0; Mismatches 57; Indels 22; Gaps 2;

QY 217 GTGAATATGAGTGGCGAGGAAGATCAACATATAGGCTAGGCCAAGAGAGAGTTTACA 276
DB 66 GTGAATATGAGTGGCGAGGAAGATCAACATATAGGCTAGGCCAAGAGAGTTTACC 125
QY 277 GCTCTGAGCTGATGGGGCTATGCTTACTGCTCCCTTTTGTCCAGAAACCCACTGA 336
DB 126 ACCTCTGAGCTGATGGGCTATGCT-----GGAGCCCGGTGA 164
QY 337 TGAAGAGCTTAAGAGAGAAACCCACCTAAAGTCGGAATCCTACACCTGATCAGAA 396
DB 165 TGAGGAGCTCAGCAAGAGAGAACCCAACTGAAAGTCGGGATCCTGACCTGGTCAGGA 224
QY 397 GAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAAAGCCGATCTCCA 456
DB 225 GAGAGAGAGATCAGGGTGCAGCTGAGACTCAAGTGCCTGACCTGGAAAGCTGATCTCCA 284
QY 457 GGAGCTATGTCAGACAAAGACTGGGGATGGATGGATGGATGGATGGATGGATGGATGGAT 516
DB 285 GGAGCTGTCTCATCAAGACTGGGGTGGATGGATGGATGGATGGATGGATGGATGGATGGAT 344
QY 517 GATTCTACCAAGCAGAGACCTTTTAAATGCCAAGCAGGTGAGGAGGAATCAGAGGT 576
DB 345 GATTCTGCCAAATCAGAACATTTTAAATGCCAAGGAGGTGACAGGCAACCCAGGT 404
QY 577 TTAAGGAGATGAGCTGAACACAC-AAAATGTTTTTATATTAGATTTTACTTTAA 635
DB 405 TTAATGAAGACAGCTGAACACACAAACAACTGTTTTTATTAAGATATTGACTTAA 464
QY 636 AGAGTCTTTAATAATTTTG 655
DB 465 AATATCAAAATAAATTTTG 484


```

RESULT 41
AI742551/c
LOCUS
DEFINITION
  AI742551 503 bp mRNA linear EST 19-DEC-1999
  w55c11.x1 Soares NSF F8 9W OT PA P S1 Homo sapiens cDNA clone
  IMAGE:2369012 3' similar to SW:GG2_HUMAN Q13066 GAGE-2 PROTEIN. [1
  //, mRNA sequence.
ACCESSION
AI742551
VERSION
AI742551.1 GI:5110839
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 503)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: 1035 Std Error: 0.00
High quality sequence stop: 454.
FEATURES
source
1..503
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2369012"
/lab_host="DH10B"
/clone_lib="Soares NSF F8 9W OT PA P S1"
/notes="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer
in a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and clones: Soares NbHSP pool 1:
30384-310919, 323208-325895 Soares Nb2HP pool 1:
145032-147335, 147220-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares Nb2HF8-9W pool 1:
758280-760583, 772104-774407 Soares NbHFA pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NbHOT
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Ronaldo."
BASE COUNT 105 a 121 c 98 g 179 t
ORIGIN
Query Match 42.0%; Score 280.8; DB 9; Length 503;
Best Local Similarity 81.1%; Pred. No. 7.6e-63;
Matches 360; Conservative 0; Mismatches 62; Indels 22; Gaps 2;

Qy 213 AATAGTCAATATACGTTGGGAGGATCAACATATAGCTAGCCAGGAGAGTT 272
Db 455 AATGTGAATATGATTGGCGAGAGATCAACATATAGCTAGCCAGGAGAGTG 396
Qy 273 TACAGCTCTGAGCTGATTGGGGCTATGCTTACTGCTCCCTTTGTCCAGGAACCCA 332
Db 395 TACCACCTCTGAGCTGATTGGGCTATGCT-----GGAGCCCG 357
Qy 333 CTGATGAAGACCTTAAGAGAGAGAACCCACCTAAAGTCGGAATCTTACCTGATC 392
Db 356 GTGATGAGGAGCCTCAGCAAGAGAGAACCAACTGAAAGTCGGATCTCTGACCTGGTC 297
Qy 393 AGAAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAAGCCGATC 452
Db 296 AGGAGAGAGAGAGATCAGGGTTCAGCTAGAGATCAAGTGCCTGACCTGGAAGCTGATC 237

```

```

Qy 453 TCCAGGAGCTATGTCAGACAAAGACTGGGATGCGATGTGAGCTGCTAGTGTCAAGG 512
Db 236 TCCAGGAGCTGTCTCAGTCAAGAGACTGGGGTGAAATGTGGAAATGGTCTCTGAGCAGG 177
Qy 513 GGAAGATTCTTACCAAAAGACAGACACTTTTAAATGCCAGAGCAGGTGAAGGAAATCAC 572
Db 176 GGAAGATTCTCCCAAAATCAGAACAAATTTAAATGCCAGAGCAGGTGACAGCAACCAC 117
Qy 573 AGTTTAAAGAGAGATGAGCTGAACACAC-AAACTGTTTTTATTAGATATTTTACT 631
Db 116 AGTTTAAATGAAGACAAAGCTGAACACCAACCAAACTGTTTTTATTAGATATTTGACT 57
Qy 632 TTAAGAGTCTTAAATAAATTTTG 655
Db 56 TAAATAATCGAATAAATTTTG 33

RESULT 42
BG436305
LOCUS
DEFINITION
  BG436305 849 bp mRNA linear EST 14-MAR-2001
  602508742F1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4605658 5',
  mRNA sequence.
ACCESSION
BG436305
VERSION
BG436305.1 GI:13342811
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 849)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLCMI343 row: d column: 11
High quality sequence stop: 514.
FEATURES
source
1..849
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4605658"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_79"
/notes="Organ: placenta; Vector: pDNR-LIB (Clontech);
Site 1: SfiI (ggcgctcgcc); Site 2: SfiI (ggcgattggcc
); 5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGCGCATATGGC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGCGGCGGACATG-dt(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."
BASE COUNT 263 a 177 c 262 g 147 t
ORIGIN
Query Match 41.6%; Score 278.6; DB 10; Length 849;
Best Local Similarity 74.2%; Pred. No. 3.4e-62;
Matches 406; Conservative 0; Mismatches 84; Indels 57; Gaps 2;

Qy 102 GACCCCTGGGAGCTGTGAGAGTGTGAGGGCAGCTTCCAGCCGCTCGGACTCTTCTCT 161

```


Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 366
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 663 Std Error: 0.00
Seq primer: M13Rp1
High quality sequence stop: 366.
Location/Qualifiers

FEATURES

source

1..544
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3885512"
/db_xref="taxon:9606"
/clone="IMAGE:252446"
/dev_stage="two placentae: one from 8 weeks and another
from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares Placenta 8to9weeks 2NHP8to9W"
/note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTACCAACTGAAGTGGGCGGCGCGATTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo."

BASE COUNT 127 a 121 c 149 g 132 t 15 others

ORIGIN

Query Match 39.9%; Score 266.8; DB 14; Length 544;
Best Local Similarity 81.3%; Pred. No. 3.7e-59;
Matches 427; Conservative 0; Mismatches 59; Indels 39; Gaps 9;
QY 31 GTTCTTGACACCTGCTCAGTGTGATGTCATCTGGGATCTTCCCTTCGACCCCTTT 90
Db 43 GCTCTTGACACCTGCTCAGTGTGATGTCATCTGGGATCTTCCCTTCGACCCCTTT 102
QY 91 GCCACGTGGTGACCGTGGGAGCTGTGAAGTGTGAGGGGACAGTTCACGCGTCTGG 150
Db 103 GCCACGTGGTGACCGTGGGAGCTGTGAAGTGTGAGGGGACAGTTCACGCGTCTGG 162
QY 151 ACTCTTCTCTCTACTGAGCGAGCTATAGTTCGCGAGCGCAGTCTCCAGGAACT 210
Db 163 ACTCTTCTCTCTACTGAGCGAGCTATAGTTCGCGA-GCCAGTCTCTCCAGGAACT 221
QY 211 GAATAGTGAATATAGTTCGCGAGGAAGATCAACATATAGGCTAGGCGGCAAGAAAG 270
Db 222 GAAATAGTGAATATAGTTCGCGAGGAAGATCAACATATAGGCTAGG-CAAGAAAG 280
QY 271 TTTACAGCTCTGAGCTGATGGGGCTATGCTTACTGCTCCCTTTGTCCCGAGAAC 330
Db 281 TTTACAGCTCTGAGCTGATGGGGCTATGCTT-----GAACT 319
QY 331 CACTGATGAAGGCTTAAAGAGAAACACCCATTAAGTCCGAA--TCCTACACCT 388
Db 320 CACTGATGAAGGCTTAAAGAGAAACACCCATTAAGTCCGAA--TCCTACACCT 379
QY 389 G-----ATCAGNAGAGAGATGATCAGGGTGCAGCTGAGATTCAA----GTGCTGACC 440
Db 380 NATTCAGNAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAAGTTCGCTGAACN 439
QY 441 TGGAGGCGGATCTCCAGGAGCTAT--GTCCAGACAAAGACTGGGATGG--ATGTGAAGG 495
Db 440 GGGAGGCGGTTTCCAGGAGGTTTGTTCAGACAAAGATGGGNTTGGATGTTGNAAGG 499
QY 496 TGGTACTGATG-TGAAGGGGAGATCTTACCAAAAGAGAGAGCACT 539
Db 500 TTGTATTGATGTTCAAGGGGAGTTTTCCTCCAAAGAGGAGGAGT 544

RESULT 45

N27023/c
LOCUS
DEFINITION
yw3el2.s1 Soares_placenta_8to9weeks_2NHP8to9W Homo sapiens cDNA
clone IMAGE:257902 3', mRNA sequence.
N27023
VERSION
N27023.1 GI:1141371
EST.
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1. (bases 1 to 430)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman,
M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
Rifkin, L., Rohlfing, T., Soares, M., Tan, P., Trevasakis, E., Waterston,
R., Williamson, A., Wohlmann, P. and Wilson, R.
The WashU-Merck EST Project
Unpublished
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 296
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: m13 -40 forward
High quality sequence stop: 296.
Location/Qualifiers

FEATURES

source

1..430
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3887512"
/db_xref="taxon:9606"
/clone="IMAGE:257902"
/dev_stage="two placentae: one from 8 weeks and another
from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares Placenta 8to9weeks 2NHP8to9W"
/note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTACCAACTGAAGTGGGCGGCGGATTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo."

BASE COUNT 88 a 115 c 85 g 141 t 1 others

ORIGIN

Query Match 39.8%; Score 266.4; DB 14; Length 430;
Best Local Similarity 81.5%; Pred. No. 4.4e-59;
Matches 353; Conservative 0; Mismatches 57; Indels 23; Gaps 3;
QY 225 TGAGTTGGCGAGAGATCAACATATAGGCTTAGCCGAGGAGAAAGATTACAGCCTCTG 284
Db 430 TGATTTGGCGAGAGATCAACATATAGGCTTAGCCGAGGAGAAAGTTTACACCTCTG 371
QY 285 AGCTGATTTGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAAACCCACTGATGAAGAC 344
Db 370 AGCTGATTTGGGCTATGCTT-----GGAGCCCGGTGATGAGGAGC 332
QY 345 CTAAAGAGAGAAACCCACC-CACTAAAGTCCGAATCTTACACCTGATCAGAGAGAGAA 403
Db 331 CTCAGCAAGNGAACCCACCAAACTGAAAGTCCGGATCTGACCTGGTCAGAGAGAGAA 272
QY 404 GATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGAGAGCCGATCTCCAGAGAGCTA 463
Db 271 GAAGATCAGGGTGCAGCTGAGACTCAAGTGCCTGAGCTGAGAGCTGATCTCCAGGAGCTG 212

```
QY 464 TGTGACAAAGACTGGGGATGGATGTGAAGTGGTACTGATGTCAAGGGGAAGATTCTA 523
Db |||||
QY 211 TCTCAGTCAAAGACTGGGGTGAATGTGGAATGGTCTCTGATGACCCAGGGGAAGATTCTG 152
Db |||||
QY 524 CCAAAAGCAGAGCACTTTTAAATGCCAGAAGCAGGTGAAGGGAATCACACGGTTTAAAGG 583
Db |||||
QY 151 CCAAAATCAGAACAATTTTAAATGCCAGAAGGAGGTGACAGGCAACCAACGGTTTAAATG 92
Db |||||
QY 584 AAGATAAGCTGAAACAACAC-AAACTGTTTTTATTATTAGATATTTTACTTTTAAAGAGTCT 642
Db |||||
QY 91 AAGACAAGCTGAACAACAACACAAAACCTGTTTTTATCTAAGATATTGACTTAAATAATCG 32
Db |||||
QY 643 TAATTAATTTTG 655
Db |||||
QY 31 AAATAAATTTTG 19
Db |||||
```

Search completed: January 18, 2004, 18:36:19
Job time : 1023.55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: April 25, 2004, 07:00:48 ; Search time 4256.62 Seconds
(without alignments)
16458.288 Million cell updates/sec

Title: US-10-051-835-25
Perfect score: 2346
Sequence: 1 gccagaggggaaaaaagag.....aaacattgatgtttacat 2346

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthm:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_estl1:*
10: gb_hcc:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gssl1:*
29: gb_gssl2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	947.8	40.4	1082	12	BM547723
C 2	742.4	31.6	762	12	BM977480
C 3	719.4	30.7	799	12	BM857800
4	692.4	29.5	806	12	BM291471

5	689	29.4	842	12	BI869486
6	688	29.3	704	12	BQ014898
7	687.6	29.1	764	12	BG177619
8	677.6	28.9	712	10	AW961203
9	665.8	28.4	737	14	CD521204
10	658	28.0	729	14	CF145565
11	642.4	27.4	644	14	CF132354
12	632	26.9	681	14	CF146273
13	625	26.6	948	10	BF796780
14	619	26.4	647	12	BM668540
15	610.8	26.0	953	12	EG254117
16	603.8	25.7	624	10	AW388283
17	597	25.4	597	14	CB138744
18	594	25.3	885	10	BE545606
19	593.6	25.3	600	10	AW388333
20	589	25.1	679	10	AW388455
21	586.2	25.0	643	12	BG944108
22	583.4	24.9	585	14	CB116564
23	580.4	24.7	608	10	AW388339
24	578.6	24.7	710	14	CD640940
25	576.8	24.6	590	12	BM740046
26	573.6	24.5	589	10	AW388414
27	573.2	24.4	631	13	BQ807293
28	573	24.4	573	14	CD677372
29	573	24.4	629	10	AW388273
30	569.2	24.3	660	10	AW378440
31	562.4	24.0	613	14	CF146275
32	562	24.0	578	13	BQ678642
33	560.2	23.9	605	10	AW388336
34	551.6	23.5	567	10	AW388480
35	544.4	23.2	588	13	BQ364149
36	544	23.2	595	10	AW388453
37	543.6	23.2	597	10	AW388607
38	540.2	23.0	608	10	AW388413
39	538.4	22.9	589	14	CF145360
40	537.2	22.9	611	10	AW388642
41	534.8	22.8	673	13	BQ364178
42	530.6	22.6	565	10	AW388711
43	529	22.5	567	10	AW388561
44	528.8	22.5	532	14	CB135713
45	526	22.4	526	10	AW237145

ALIGNMENTS

RESULT 1
BM547723
LOCUS
DEFINITION
AGENCOURT_6507143 NTH_MGC_124 Homo sapiens cDNA clone IMAGE:5727848
5', mRNA sequence.
ACCESSION
BM547723
VERSION
BM547723.1 GI:18781736
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. (bases 1 to 1082)
NIH-MGC <http://imgc.ncbi.nlm.nih.gov/>.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12722 row: j column: 09
High quality sequence start: 3

High quality sequence stop: 723.

FEATURES

Location/Qualifiers
1..1082
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5727848"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH MGC 124"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: EcoRV (destroyed); Site 2: NotI; RNA source male hippocampus, age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 0.9-4 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 012."

ORIGIN

Query Match 40.4%; Score 947.8; DB 12; Length 1082;

Best Local Similarity 96.4%; Pred. No. 3.5e-180; Mismatches 31; Indels 7; Gaps 6;
Matches 1031; Conservative 0;

QY	378	GAACAGATTGTTCCCATGAATGTATGCTCATATCTGACTAGAAACGATCTATGTTAAAT	437
Db	16	GATACAGATTGTTCCCATGAATGTATGCTCATATCTGACTAGAAACGATCTATGTTAAAT	75
QY	438	GACTGTGATATGAATATTTTCAAGTACTACCCCAATTAACCTTCTTATGCTCTGAAAG	497
Db	76	GACTGTGATATGAATATTTTCAAGTACTACCCCAATTAACCTTCTTATGCTCTGAAAG	135
QY	498	AAGAAAGCAATGAATCACTATGATTTATGCAACAAACAGAAATCTCCAAAT	557
Db	136	AAGAAAGCAATGAATCACTATGATTTATGCAACAAACAGAAATCTCCAAAT	195
QY	558	TTAAGTAATCTGCTCTCTCTGAGAAATGTTACCTAATAGTATTTTCTTATGAT	617
Db	196	TTAAGTAATCTGCTCTCTCTGAGAAATGTTACCTAATAGTATTTTCTTATGAT	255
QY	618	GTTATTACTACTGGTATAATCAATCTTCTATAAATTTCTACTTAAAGTCCTTAAGAACTG	677
Db	256	GTTATTACTACTGGTATAATCAATCTTCTATAAATTTCTACTTAAAGTCCTTAAGAACTG	315
QY	678	GGTCTTCTCTTGAATGATATGCTGTCAGAAAGGAAACACACATTTACTCTTTAGGAC	737
Db	316	GGTCTTCTCTTGAATGATATGCTGTCAGAAAGGAAACACACATTTACTCTTTAGGAC	375
QY	738	AATTCTCTAGAACTCTATAGTAGTATCAGGATATATTTTGTCTTAAATATATTTTGGTTAT	797
Db	376	AATTCTCTAGAACTCTATAGTAGTATCAGGATATATTTTGTCTTAAATATATTTTGGTTAT	435
QY	798	TTTGAATACAGACATTGGCTCCAAATTTTCATCTTTGCAATAGTATGACTTTTCACTA	857
Db	436	TTTGAATACAGACATTGGCTCCAAATTTTCATCTTTGCAATAGTATGACTTTTCACTA	495
QY	858	GAATCTCTCAACATTTGGGAACTTTGCAATATGAGCATCATATGTTGTTAAGCTGTATC	917
Db	496	GAATCTCTCAACATTTGGGAACTTTGCAATATGAGCATCATATGTTGTTAAGCTGTATC	555
QY	918	ATTTAATGCTATGAGATACATTTGTTTCTCCCTATGCAACAGGTGAACAACTAGTIT	977
Db	556	ATTTAATGCTATGAGATACATTTGTTTCTCCCTATGCAACAGGTGAACAACTAGTIT	615
QY	978	GTTTTTTTACTGATACTAAATGTTGGCTACCTGTTGATTTTATAGTATGCACTGAGAAA	1037
Db	616	GTTTTTTTACTGATACTAAATGTTGGCTACCTGTTGATTTTATAGTATGCACTGAGAAA	675
QY	1038	AAGCAAGACAAATGGCTCTTGTACTGAATATCTCGGCAAACTTATTTGGGTCTTCATTT	1097
Db	676	AAGCAAGACAAATGGCTCTTGTACTGAATATCTCGGCAAACTTATTTGGGTCTTCATTT	735
QY	1098	TCTGACAGACAGGATTTGACTCAATA-TTTGTAGAGCTTCGGTAGAATGATACATGGT	1156

Db	736	TCTGACAGACAGGATTTGACTCAATATTTTGTAGAGCTTCGGTAGAATGATACATGT	795
QY	1157	AGTGATGCACTGGTAGAAATGTTTTTATGTTATGACTCAGAAATTCATCTCAGGATGAAT	1216
Db	796	AGTGATGCACTGGTAGAAATGTTTTTATGTTATGACTCAGAAATTCATCTCAGGATGAAT	855
QY	1217	CTTTTATGCTCTTTTATGTTATGTAAGCATATCTGAATTTTACTTTTATAAGATGGTTTTAGAAA	1276
Db	856	CTTTTATGCTCTTTTATGTTATGTAAGCATATCTGAATTTTACTTTTATAAGATGGTTTTAGAAA	914
QY	1277	GCTTTCTCTAAAAATTTGGCTAGGAATGGTAACCTTCATTTTCAGTTCGCAAGGGGTAGA	1336
Db	915	GCTTTCTCTAAAAA-TTGGCTAGGAATGGTAACCTTCATTTTCAGTTCGCAAGGGGTAGA	973
QY	1337	AAATATATATGCTGTGTTGTTAT-GTTTATGTTTAAACATATATTA-GGTACTATCTATGAA	1394
Db	974	AAATATATATGCTGTGTTGTTATGTTTAACTTATATAGGTTACTATCTATGAA	1033
QY	1395	TGTATTTAAATA--TTTTTCATATCTCTGACAAAGCAATTTTAAATTTGC	1441
Db	1034	AGTAATTAATAATTTTCTAAATCTGGGAAACCAATTTTAAATTTGC	1082

RESULT 2

BM977480/c
LOCUS
DEFINITION
UI-CF-EN1-ase-b-22-0-UI-s1 UI-CF-EN1 Homo sapiens cDNA clone
UI-CF-EN1-ase-b-22-0-UI 3', mRNA sequence.

ACCESSION
BM977480VERSION
BM977480.1 GI:19595939KEYWORDS
EST.SOURCE
Homo sapiens (human)ORGANISM
Homo sapiensEukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.REFERENCE
1 (bases 1 to 762)AUTHORS
Bonaldo, M.F., Lennon, G. and Soares, M.B.TITLE
Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

MEDLINE
97044477PUBMED
8889548COMMENT
Contact: McCray, PB

McCrack Lab

University of Iowa

2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866

Fax: 319 356 7171

Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com) or from Open Biosystems

(www.openbiosystems.com).

The following repetitive elements were found in this cDNA

sequence: 1-64, >AT rich#Low_complexity

Seq primer: M13 FORWARD

POLYA=Yes.

FEATURES

source

Location/Qualifiers

1..762

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clones="UI-CF-EN1-ase-b-22-0-UI"

/tissue_type="Primary Lung Cystic Fibrosis Epithelial

Cells"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone_lib="UI-CF-EN1"

/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-EN1 is a normalized cDNA library containing the following tissue(s): Primary Lung Cystic Fibrosis Epithelial Cells. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CTGCTCAGGT. TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS 6hr to LPS 24h TAG LIB=UI-CF-EN1 TAG_SEQ=CTGCTCAGGT"

ORIGIN

Query Match 31.6%; Score 742.4; DB 12; Length 762;
Best Local Similarity 99.9%; Pred. No. 6.5e-139;
Matches 743; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1603 CATTCGCCAGATTTCAGGGGATCTGGGTTTGGTCTAGTGAATGAAACACAAAGACAG 1662
DB 762 CATTCGCCAGATTTCAGGGGATCTGGGTTTGGTCTAGTGAATGAAACACAAAGACAG 703
QY 1663 AGAGATCCAGCTGAAAAGAGTGATCCTCAATATCCTAACTAACTGTCCTCAACTCAAG 1722
DB 702 AGAGATCCAGCTGAAAAGAGTGATCCTCAATATCCTAACTAACTGTCCTCAACTCAAG 643
QY 1723 CAGAGTTTCTTCACCTCTGCACTGTGATCATGAACTTTAGTAGAGGGGATGTGTGATT 1782
DB 642 CAGAGTTTCTTCACCTCTGCACTGTGATCATGAACTTTAGTAGAGGGGATGTGTGATT 583
QY 1783 TTATACAAATTAAATACAAATGCTTCAATGATGATAAAATCTTAAGACAAACTGCATT 1842
DB 582 TTATACAAATTAAATACAAATGCTTCAATGATGATAAAATCTTAAGACAAACTGCATT 523
QY 1843 TTATTTCTGATCCATCCATCCATATATGAACTAAGATATTTATCTATGAAGATATA 1902
DB 522 TTATTTCTGATCCATCCATCCATATATGAACTAAGATATTTATCTATGAAGATATA 463
QY 1903 AATGTGTCAGAGAGACTTTCATCTGTGGAATGCGTGTGTTCTTAGGGTTCCTAGCACTGA 1962
DB 462 AATGTGTCAGAGAGACTTTCATCTGTGGAATGCGTGTGTTCTTAGGGTTCCTAGCACTGA 403
QY 1963 TGCCTGCAACAGCATGTGATATGGAATAAATGATTTCTTCTATAGTAAATCAGTTC 2022
DB 402 TGCCTGCAACAGCATGTGATATGGAATAAATGATTTCTTCTATAGTAAATCAGTTC 343
QY 2023 CCTCTGGGAGAGTTCCTGCTACTGCAATCACAATGCCAGATGCTGTTTATGGGCTATTG 2082
DB 342 CCTCTGGGAGAGTTCCTGCTACTGCAATCACAATGCCAGATGCTGTTTATGGGCTATTG 283
QY 2083 TGTAGTAAAGTAAAGTCTATGAGTAAGTGTGTTTGTTCATCTTATGGAACATC 2142
DB 282 TGTAGTAAAGTAAAGTCTATGAGTAAGTGTGTTTGTTCATCTTATGGAACATC 223
QY 2143 TTGATGCATGTGTTTTGTATGGAATAAATTTTGGTGCAATATGATGTCATTCAACTTTG 2202
DB 222 TTGATGCATGTGTTTTGTATGGAATAAATTTTGGTGCAATATGATGTCATTCAACTTTG 163
QY 2203 CATTGAATTGAATTTGGTGTATTTATGATATATATCTGCTCAGGCTCTAGTGTGTT 2262
DB 162 CATTGAATTGAATTTGGTGTATTTATGATATATATCTGCTCAGGCTCTAGTGTGTT 103
QY 2263 CAACCAATTTATACCAATTTTGTACATATTTTACTTGAATAATTTTAAATGGAATTT 2322
DB 102 CAACCAATTTTATACCAATTTTGTACATATTTTACTTGAATAATTTTAAATGGAATTT 43
QY 2323 AAATAAACATTTGATGTTTACAT 2346

Db 42 AAATAAACATTTGTTAGTTTACAT 19

RESULT 3
BI857800/c
LOCUS BI857800.1
DEFINITION 799 bp mRNA linear EST 10-OCT-2001
60338850F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5397469 5',
mRNA sequence.
ACCESSION BI857800
VERSION BI857800.1 GI:15998547
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 799)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: DCTD/DMP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL2013 row: d column: 14
High quality sequence stop: 776.
Location/Qualifiers
FEATURES
source
1..799
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5397469"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="NIH_MGC_87"
/clone_lib="NIH_MGC_87"
/notes="Organ: breast; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 30.7%; Score 719.4; DB 12; Length 799;
Best Local Similarity 97.4%; Pred. No. 2.7e-134;
Matches 753; Conservative 0; Mismatches 16; Indels 4; Gaps 2;

QY 1571 GGATCTGGATGATGATTTATGACCTTGTACCTTCCAGATTTCCAGGGGATCGTGG 1630
DB 799 GGATCTGGATGATGATTTATGACCTTGTACCTTCCAGATTTCCAGGGGATCGTGG 740
QY 1631 GTTTGGTCTAGTGAATGAAACACAAAGACAGAGATCC---AGCTGAAAAAGATGAT 1687
DB 739 GTTTGGTCTAGTGAATGAAACACACTAGTCCAGAGAGATTCAGCTTGAAAAAGAGATGAT 680
QY 1688 CCTCAATATCCTTAACCTGCTCCCTCACTCAAGCAGAGTTTCTTCACTCTGGCACTGT 1747
DB 679 CCTCAATATCCTTAACCTGCTCCCTCACTCAAGCAGAGTTTCTTCACTCTGGCACTGT 620
QY 1748 GATCATGAAACTTAGTAGAGGGGATGTGTGTTATTTTATACAAATTTAATACATGTCCT 1807
DB 619 GATCATGAAACTTAGTAGAGGGGATGTGTGTTATTTATACAAATGTTAATACATGTCCT 560
QY 1808 ACATTGATAAATTTCTTAAAGAGCAAACTGCATTTA-TTTCTGATCCACATTCCTCAAT 1866
DB 559 ACATTGATAAATTTCTTAAAGAGCAAACTGCATTGAGTTTCTGCAATCCATTCCTCAAT 500
QY 1867 CATATTGAACTAGATATTTATCTATGAAGATATAAATGTTGTCAGAGAGACTTTTCACTCT 1926
DB 499 CATATTGAACTAGATATTTATCTATGAAGATATAAATGTTGTCAGAGAGACTTTTCACTCT 440


```

Db      600  CGTGTTCCTTAGGTTCTAGCACTGATGCTGCACAGCATGTGATATGTGAATAACA 659
Qy      1995  ATGATTTCTTTATAGCTAAATAGTTCCTCTGCGGAGAGTTCTGTGACTGCAATCACA 2054
Db      660  ATGATTTCTTTATAGCTAAATAGTTCCTCTGCGGAGAGTTCTGTGACTGCAATCACA 719
Qy      2055  ATGCCAGATGGTGTATTATGGGCTAT 2079
Db      720  ATGCCAATGGTGTATTATGGGCAAT 744

RESULT 8
LOCUS      AW961203              712 bp  mRNA  linear  EST 01-JUN-2000
DEFINITION EST373275 MAGE resequences, MAGEF Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW961203
VERSION    AW961203.1 GI:8150887
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 712)
AUTHORS   Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspar, R., Gay, C.,
            Holt, I. E., Saeed, A. I., Sharov, V., Lee, N. H., Yeatman, T. J. and
            Quackenbush, J.
TITLE     Assessment of gene expression patterns in a model of colon tumor
          metastasis using a 19,200 element cDNA microarray
JOURNAL   Unpublished (2000)
COMMENT   Contact: John Quackenbush
          The Institute for Genomic Research
          9712 Medical Center Dr., Rockville, MD 20850, USA
          Tel: 301 838 3528
          Fax: 301 838 0208
          Email: johnq@tigr.org
          Plate: 155
          Seq primer: Reverse.
FEATURES  Location/Qualifiers
            source          1..712
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone_lib="MAGE resequences, MAGEF"
                        /note="Vector: pBluescriptSKm"
ORIGIN
Query Match      28.9%; Score 677.6; DB 10; Length 712;
Best Local Similarity 99.2%; Pred. No. 6.7e-126;
Matches 702; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
Qy      29  AGGTATCTCTTTGCGAGTGGTCACTGTATTTTGTAGTACCTTCTGTGACAGGGTATTATT 88
Db      1  AGGTATCTCTTTGCGAGTGGTCACTGTATTTTGTAGTACCTTCTGTGACAGGGTATTATT 59
Qy      89  ACAGCATCTTGTGGGAAACCTATTAGGCTTTGATGCTTTAAGTGTATATTGTTGG 148
Db      60  ACAGCATCTTGTGGGAAACCTATTAGGCTTTGATGCTTTAAGTGTATATTGTTGG 119
Qy      149  GTTGTGAGTGGTCTGACTTAATGTATATTAATAATTTAGACATCAAAATTTTCTACTA 208
Db      120  GTTGTGAGTGGTCTGACTTAATGTATATTAATAATTTAGACATCAAAATTTTCTACTA 179
Qy      209  ACTAATCTTTTAGATGCACTATTGGAAGCAGAGTCATATCATCTGGGAGCAATGCAA 268
Db      180  ACTAATCTTTTAGATGCACTATTGGAAGCAGAGTCATATCATCTGGGAGCAATGCAA 239
Qy      269  TGTGGTTACTGCTGCTAGTTTGAAGTCTTATTTCGAAGATTTCTGAAATTAATTT 328
Db      240  TGTGGTTACTGCTGCTAGTTTGAAGTCTTATTTCGAAGATTTCTGAAATTAATTT 299
Qy      329  TCCTTAGAATTTCTCTTCATTCCTCAAGTACAAACATCTTTTGAAGATGAAACAGATTG 388
Db      300  TCCTTAGAATTTCTCTTCATTCCTCAAGTACAAACATCTTTTGAAGATGAAACAGATTG 359

```

```

Qy      389  TTCCCATGATGATGCTCATACTCGACTAGAAAAGATCTATGTTAAATGACTGTGTATA 448
Db      360  TTCCCATGATGATGCTCATACTCGACTAGAAAAGATCTATGTTAAATGACTGTGTATA 419
Qy      449  TGAATTTATTTCAGTACTACCCCAATAACTTTCTTTATGCTCTGAAAGAGAAAAGCAA 508
Db      420  TGAATTTATTTCAGTACTACCCCAATAACTTTCTTTATGCTCTGAAAGAGAAAAGCAA 479
Qy      509  TGTAAATCACTATGATTATTGCACAAACACAGAAATCTCCCAACAATTTTAAAGTAATCT 568
Db      480  TGTAAATCACTATGATTATTGCACAAACACAGAAATCTCCCAACAATTTTAAAGTAATCT 539
Qy      569  GATCCCTCTTCTTGGAGAAAATTTGTTACCTTAATAGTTTTTCCCTTATGAATGTTATTACTAC 628
Db      540  GATCCCTCTTCTTGGAGAAAATTTGTTACCTTAATAGTTTTTCCCTTATGAATGTTATTACTAC 599
Qy      629  TGGTATAAATCAAAATTTCTATAAATTTCTCTTAAGTCTTAAAGAACTGGGTTCTTCCCTT 688
Db      600  TGGTATAAATCAAAATTTCTATAAATTTCTCTTAAGTCTTAAAGAACTGGGTTCTTCCCTT 659
Qy      689  TGATGTTATTTCATGTTTCAGAAA-GGAAACAACACTTTACTCTTTTAGG 735
Db      660  TGATGTTATTTCATGTTTCAGAAAAGGAGAAACAACACTTTACTTTTAGG 707

RESULT 9
CD521204              737 bp  mRNA  linear  EST 06-JUN-2003
LOCUS      AGENCOURT 14353302 NIH MGC 191 Homo sapiens cDNA clone
DEFINITION IMAGE:30412131 5', mRNA sequence.
ACCESSION  CD521204
VERSION    CD521204.1 GI:31452922
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 737)
AUTHORS   NIH-MGC http://mgs.nci.nih.gov/
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgabs-x@mail.nih.gov
            Tissue Procurement: Narayan Bhat
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
Plate: NDCM205 row: a column: 04
High quality sequence stop: 592.
FEATURES  Location/Qualifiers
            source          1..737
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone_lib="IMAGE:30412131"
                        /tissue_type="Pooled"
                        /lab_host="DH10B (T1 phage-resistant)"
                        /clone_lib="NIH MGC 191"
            Note: Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);
            Site 2: SfiI (ggcgctctggcc); Library is oligo-dT primed
            and directionally cloned. pBMC - Peripheral Blood
            Mononuclear Cells. RNA was pooled from 3/6hour stimulation
            with PMA adn Ionomycin. 5' and 3' adaptors were used in
            cloning as follows: 5' adaptor sequence:
            5'-CACGGCCATTATGCC-3' and 3' adaptor sequence:
            5'-ATTCTAGGCGCGAGCGCCGACATG-DT(30)BN-3' (where B = A,

```

C, or G and N = A, C, G, or T). Average insert size 1.69 kb (range 0.70-5.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN

```
Query Match      28.4%; Score 665.8; DB 14; Length 737;
Best Local Similarity 99.7%; Pred. No. 1.6e-123;
Matches 667; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1033 AGAAAAAGGCAAGCAAAATGGCCCTCTGTACTGAATACCTTCGGCAAACTTATTTGGTCTT 1092
DB 3 AGAARAGGCAAGCAAAATGGCCCTCTGTACTGAATACCTTCGGCAAACTTATTTGGTCTT 62
QY 1093 CATTTCTGACAGACAGATTTGACTCAATATTGTAGACCTTGGCGTAGAATGATACA 1152
DB 63 CATTTCTGACAGACAGATTTGACTCAATATTGTAGACCTTGGCGTAGAATGATACA 122
QY 1153 TGGTAGTGATGCACTGGTAGAAATGGTTTTAGTTATTGACTCAGAAATTCATCTCAGGAT 1212
DB 123 TGGTAGTGATGCACTGGTAGAAATGGTTTTAGTTATTGACTCAGAAATTCATCTCAGGAT 182
QY 1213 GAATCTTTATGCTCTTTTATTTGTAAGCATATCTGAATTTACTTTTATAAGATGTTTTA 1272
DB 183 GAATCTTTTATGCTCTTTTATTTGTAAGCATATCTGAATTTACTTTTATAAGATGTTTTA 242
QY 1273 GAAAGCTTTGTCTAAAAATTTGGCTAGGAATGTAATCTTCAATTTTCAGTTGCGCAAGGGG 1332
DB 243 GAAAGCTTTGTCTAAAAATTTGGCTAGGAATGTAATCTTCAATTTTCAGTTGCGCAAGGGG 302
QY 1333 TAGAAAAATATATGCTGTGTTGTTATGTTTATGTTATGTTATTAATTTAGTACTATCTATG 1392
DB 303 TAGAAAAATATATGCTGTGTTGTTATGTTTATGTTATTAATTTAGTACTATCTATG 362
QY 1393 AATGTTATTAATATTTTTCATATCTGTGACAGCATTTTATATTTTGCAACAGTGGAG 1452
DB 363 AATGTTATTAATATTTTTCATATCTGTGACAGCATTTTATATTTTGCAACAGTGGAG 422
QY 1453 TCCATTTAGCCAGTGGGAAAGCTTTGGAATCTAGTTACCTTTCCAGGAATTTTCAGGGGATG 1512
DB 423 TCCATTTAGCCAGTGGGAAAGCTTTGGAATCTAGTTACCTTTCCAGGAATTTTCAGGGGATG 482
QY 1513 CCCATCTTTTATGCTGTGCTTAACTGTAATTTTATAGACAGCTAAATCCCTAACTTGG 1572
DB 483 CCCATCTTTTATGCTGTGCTTAACTGTAATTTTATAGACAGCTAAATCCCTAACTTGG 542
QY 1573 ATCTGGAATGATTAGTTATGACCTTTGACCTTTCCAGGAATTTTCAGGGGATGTTGGGT 1632
DB 543 ATCTGGAATGATTAGTTATGACCTTTGACCTTTCCAGGAATTTTCAGGGGATGTTGGGT 602
QY 1633 TTGCTCTAGTATGAAGAACACAGACAGAGAGATCCAGCTGAAAGAGATGATCCTCA 1692
DB 603 TTGCTCTAGTATGAAGAACACAGACAGAGAGATCCAGCTGAAAGAGATGATCCTCA 662
QY 1693 ATATCCTAA 1701
DB 663 ATATCCTAA 671
```

```
RESULT 10
CF145565/c
LOCUS      729 bp      mRNA      linear      EST 06-AUG-2003
DEFINITION UI-HF-CB0-asq-e-04-0-UI.r1 NIH_MGC_210 Homo sapiens CDNA clone
IMAGE:30569331 5', mRNA sequence.
```

```
ACCESSION  CF145565
VERSION     CF145565.1  GI:33261009
```

KEYWORDS

EST.

Homo sapiens (human)

ORGANISM

EuKaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 729)

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Tim Ratliff
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/humanfl.html
The following repetitive elements were found in this cDNA
sequence: 27-49, >POLY_A#simple_repeat (matched complement)
Seq primer: PYX-5.

FEATURES

Location/Qualifiers

1..729

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30569331"

/tissue_type="CNCAP(3)T-225 cell line"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH_MGC_210"

/note="Organ: Prostate; Vector: p7T3 Pac; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into p7T3 Pac vector. The library tag sequence located between the Not I site and the polyA tail is CCCAC. Tissue was provided by Tim Ratliff."

ORIGIN

```
Query Match      28.0%; Score 658; DB 14; Length 729;
Best Local Similarity 99.7%; Pred. No. 5.8e-122;
Matches 680; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1665 AGATCCAGCTGAAAAGAGTGAATCTCAATATCTTAATCTAATCTGCTCTCACTCAAGCA 1724
DB 729 AGATCCAGCTGAAAAGAGT-ATCCTCAATATCTTAATCTAATCTGCTCTCACTCAAGCA 672
QY 1725 GAGTTTCTTCTACTGCGCACTGTGATCATGAACTTAGTAGGGGATTTGTGTATTTT 1784
DB 671 GAGTTTCTTCTACTGCGCACTGTGATCATGAACTTAGTAGGGGATTTGTGTATTTT 612
QY 1785 ATACAAATTAATAAATATGCTTACATTTGATTAATTTCTTAAGAGCAAACTGCAATTT 1844
DB 611 ATACAAATTAATAAATATGCTTACATTTGATTAATTTCTTAAGAGCAAACTGCAATTT 552
QY 1845 ATTTCTGCATCCACATTCCTCAATCATATTAGAACTAAGATATTTATCTATGAAGATATAA 1904
DB 551 ATTTCTGCATCCACATTCCTCAATCATATTAGAACTAAGATATTTATCTATGAAGATATAA 492
QY 1905 TGGTCAGAGAGACTTTCATCTGTGGATTTGCGTTCTTTCTAGGGTTCTTACACTGATG 1964
DB 491 TGGTCAGAGAGACTTTCATCTGTGGATTTGCGTTCTTTCTAGGGTTCTTACACTGATG 432
QY 1965 CTTGCAACAGCATGATGATATGTAATAAATGGAATTTCTTATAGTAAATAGTTCCTCC 2024
DB 431 CTTGCAACAGCATGATGATATGTAATAAATGGAATTTCTTATAGTAAATAGTTCCTCC 372
QY 2025 TCTGGGGAGAGTCTGGTACTGCAATCAATGCCAGATGGTGTATGGGCTATTTGTG 2084
```


cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/humanfl.html>
The following repetitive elements were found in this cDNA
sequence: 27-49, >P0U1_AHSimple_repeat (matched complement)
Seq primer: pYX-5.

FEATURES

Location/Qualifiers

1..681
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30569820"
/tissue_type="CNAP(3)T-225 cell line"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH MGC 210"
/notes="Organ: Prostate; Vector: pTV73 Pac; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with Not I and then cloned
directionally into pTV73 Pac vector. The library tag
sequence located between the Not I site and the polyA tail
is CCCAC. Tissue was provided by Tim Ratliff."

ORIGIN

Query Match 26.9%; Score 632; DB 14; Length 681;
Best Local Similarity 100.0%; Pred. No. 9.7e-117;
Matches 632; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1715 AACTCAGCAGAGTTCTTCCACTGCGACATCATCAAACTTAGTAGAGGGATTG 1774
DB 681 AACTCAGCAGAGTTCTTCCACTGCGACATCATCAAACTTAGTAGAGGGATTG 622

QY 1775 TGTGTATTTTATACAAATTTAATACAAATGCTTACATGATGATAAAATCTTAAAGAGCAAA 1834
DB 621 TGTGTATTTTATACAAATTTAATACAAATGCTTACATGATGATAAAATCTTAAAGAGCAAA 562

QY 1835 ACTGCATTTTATCTGCATCCCATTCATCATATATAGACTAGATATTTATCTATG 1894
DB 561 ACTGCATTTTATCTGCATCCCATTCATCATATATAGACTAGATATTTATCTATG 502

QY 1895 AAGATATAAATGGTGCAGAGAGCTTTTCATCTGCGAATGCGTTGTTTCTAGGGTTCCT 1954
DB 501 AAGATATAAATGGTGCAGAGAGCTTTTCATCTGCGAATGCGTTGTTTCTAGGGTTCCT 442

QY 1955 AGCACTGATGCTGCAGACAGCATGTATGTGAATAAATGATTTCTTCTATAGCTAA 2014
DB 441 AGCACTGATGCTGCAGACAGCATGTATGTGAATAAATGATTTCTTCTATAGCTAA 382

QY 2015 ATGAGTTCCTCTGGGAGAGTTCTGCTACTGCAATCACAATGCCAGATGGTGTATGG 2074
DB 381 ATGAGTTCCTCTGGGAGAGTTCTGCTACTGCAATCACAATGCCAGATGGTGTATGG 322

QY 2075 GCTATTTGTGTAAGTAACTGATGCTATGAGTAAGTGTGTTGTTTCTATCTTAT 2134
DB 321 GCTATTTGTGTAAGTAACTGATGCTATGAGTAAGTGTGTTGTTTCTATCTTAT 262

QY 2135 GGAACTCTTGATGATGCTGTTTGTATGGAATAAATTTGGTGCATATGATGTCAAT 2194
DB 261 GGAACTCTTGATGATGCTGTTTGTATGGAATAAATTTGGTGCATATGATGTCAAT 202

QY 2195 CAACCTTCGATGAATGAATTTGGTGTGATTTATATGATATATACCTGCAGCTTCT 2254
DB 201 CAACCTTCGATGAATGAATTTGGTGTGATTTATATGATATATACCTGCAGCTTCT 142

QY 2255 AGTTCCTTCAACCAATTTTATACCAATTTTGTACATATTTTACTTGAATAATTTTAAAT 2314
DB 141 AGTTCCTTCAACCAATTTTATACCAATTTTGTACATATTTTACTTGAATAATTTTAAAT 82

QY 2315 GGAAATTTAAATAAACAATTGATGACTTTACAT 2346
DB 81 GGAAATTTAAATAAACAATTGATGACTTTACAT 50

RESULT 13

BF796780 948 bp mRNA linear EST 12-JAN-2001
LOCUS 602259726F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4342974 5',
DEFINITION mRNA sequence.

ACCESSION BF796780

VERSION BF796780.1 GI:12101834

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 948)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cchapb-remail.nih.gov

Tissue Procurement: Louis Staudt, M.D., Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LHM9559 row: C column: 07

High quality sequence stop: 843.

Location/Qualifiers

FEATURES

source

1..948

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4342974"

/tissue_type="lymphoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_85"

/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 1.867 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

ORIGIN

Query Match 26.6%; Score 625; DB 10; Length 948;
Best Local Similarity 97.9%; Pred. No. 2.4e-115;
Matches 655; Conservative 0; Mismatches 10; Indels 4; Gaps 2;

QY 14 AAAGAGTAATGCACAGGATCTCTTTTGCAGTGGTACTGATTTTTCAGTACCTTTTG 73
DB 1 AAAGAGTAATGCACAGGATCTCTTTTGCAGTGGTACTGATTTTTCAGTACCTTTG 59

QY 74 TCACAGGTAATTTATTTACAGCATCTTTGGGAAAACCTATTAGGCTTTGCATGTTAAGC 133
DB 60 TCACAGGTAATTTATTTACAGCATCTTTGGGAAAACCTATTAGGCTTTGCATGTTAAGC 119

QY 134 TGTATAATTTGTTGGTGTGAGTGGTCTGACTTAATGCTGATTAATAAATTTAGACAT 193
DB 120 TGTATAATTTGTTGGTGTGAGTGGTCTGACTTAATGCTGATTAATAAATTTAGACAT 179

QY 194 CAAATTTTCTACTAACTAACTTTTATTAGATGACATCTTGGAGCACAGTCAATATCAAC 253
DB 180 CAAATTTTCTACTAACTAACTTTTATTAGATGACATCTTGGAGCACAGTCAATATCAAC 239

QY 254 TGGAGGCAATGCATGTTGTTTACCTGCTTACCTGCTTACCTGCTTACCTGCTTATTTTCAAAAGAT 313
DB 240 TGGAGGCAATGCATGTTGTTTACCTGCTTACCTGCTTACCTGCTTATTTTCAAAAGAT 299

QY 314 TTCTGAATTAATTTTCCCTAGAAATTTCTCTTCATTCCTCCAAAGTACAAACATCATTTTGA 373

LOCUS BG254117 953 bp mRNA linear EST 13-FEB-2001
DEFINITION 602367005F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4475233 5',
mRNA sequence.
ACCESSION BG254117
VERSION BG254117.1 GI:12763933
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 953)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10300 row: j column: 02
High quality sequence stop: 642.
FEATURES
Location/Qualifiers
1..953
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4475233"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_91"
/note="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
ORIGIN
Query Match 26.0%; Score 610.8; DB 12; Length 953;
Best Local Similarity 92.1%; Pred. No. 1.7e-112;
Matches 722; Conservative 0; Mismatches 52; Indels 10; Gaps 7;
QY 494 AAGAAGAAAAGCAATGTAATCACTATGATTATGACAAACCAACAGAAATCTCCAAC 553
DB 1 AAGAAGAAAAGCAATGTAATCACTATGATTATGACAAACCAACAGAAATCTCCAAC 59
QY 554 AATTTTAAGTAATCTGATCCCTCTCTTGGAGAAAATGTTACCTAATAGTTTTCCTTAT 613
DB 60 AATTTTAAGTAATCTGATCCCTCTCTTGGAGAAAATGTTACCTAATAGTTTTCCTTAT 118
QY 614 GAATGTTATTACTACTCGTATAATCAAAATTCCTATAAATTCCTACTTAAGTCTTAAGA 673
DB 119 GAATGTTATTACTACTCGTATAATCAAAATTCCTATAAATTCCTACTTAAGTCTTAAGA 178
QY 674 ACTGGGTTCTCTTTCATGTTATTCATGTTTCAGAAAGGAAACAAACATTTACTCTTTTA 733
DB 179 ACTGGGTTCTCTTTCATGTTATTCATGTTTCAGAAAGGAAACAAACATTTACTCTTTTA 238
QY 734 GGCAATTCCTAGTAATCTAGTAGTATCAGGATATATTTGCTTTAAATATATTTTGG 793
DB 239 GGCAATTCCTAGTAATCTAGTAGTATCAGGATATATTTGCTTTAAATATATTTTGG 295
QY 794 TTATTTTGAATACAGACATTTGGCTCCAAATTTTCACTTTTGCACAAATAGTATGACATTTTC 853
DB 296 GTTATTTTGAATACAGACATTTGGCTCCAAATTTTCACTTTTGCACAAATAGTATGACATTTTC 355
QY 854 ACTAGAACTTCTCAACATTTGGGAACCTTCGAATATGAGCATCATATGTTAAGGCTG 913
DB 356 ACTAGAACTTCTCAACATTTGGGAACCTTCGAATATGAGCATCATATGTTAAGGCTG 415

QY 914 TATCAATTAATGCTATGAGATACATTTGTTTCTCCCTATGCCAAACAGGTGAACAAACGT 973
DB 416 TATCAATTAATGCTATGAGATACATTTGTTTCTCCCTATGCCAAACAGGTGAACAAACGT 475
QY 974 AGTTGTTTTTACTGATCTAATAATGTTGGCTACCTGCTGATTTTATAGTATGCACATGTCA 1033
DB 476 AGTTGTTTTTACTGATCTAATAATGTTGGCTACCTGCTGATTTTATAGTATGCACATGTCA 535
QY 1034 GAAAAAGGCAAGCAAAATGGCCCTCTTGTACTGAATACTTTGGCAAACTTATTTGGGTCTTC 1093
DB 536 GAAAAAGGCAAGCAAAATGGCCCTCTTGTACTGAATACTTTGGCAAACTTATTTGGGTCTTC 595
QY 1094 ATTTTCTGACAGA-CAGGATTTGACTCAATATTTGTAGAGCTTGCCTAGATCGATTACA 1152
DB 596 ATTTTCTGACAGACCAGGATTTGACTCAATATTTGTAGAGCTTGCCTAGATCGATTACA 655
QY 1153 TGGTAGTGTGATGCTAGTAAATGGTTTGTAGTTATTTGACTCAGAAATTCATCTCAGGAT 1212
DB 656 TGG-AATGATGCCCTGGTAGAAA--GGTCAAGTTATAGACTCCGAA-TCATCTCAGGAG 711
QY 1213 GAATCTTTTATGCTCTTTTATTTGTAAGCATATCTGAAATTTACTTTTAAAGATGGTTTAA 1272
DB 712 AAATCTTTATGCTCTCTATGGAAGGCCCTTCTGATTACTAATAACGATGGTTAAAGCTTGC 771
QY 1273 GAAA 1276
DB 772 TAAA 775

Search completed: April 26, 2004, 01:59:42
Job time : 4260.62 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 25, 2004, 07:15:18 ; Search time 116.382 Seconds
(without alignments)
11186.557 Million cell updates/sec

Title: US-10-051-835-25
Perfect score: 2346
Sequence: 1 gccagaggggaaaaaagag.....aaacatttgatgattacat 2346

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	581.4	24.8	675	4	US-09-702-705-1269	Sequence 1269, Ap
2	581.4	24.8	675	4	US-09-736-457-1269	Sequence 1269, Ap
3	581.4	24.8	675	4	US-09-671-325-1269	Sequence 1269, Ap
4	551.4	23.5	665	4	US-09-614-1248-1269	Sequence 1269, Ap
C 5	54.6	2.3	640681	4	US-09-790-988-1	Sequence 1, Appl
C 6	49.8	2.1	1501	4	US-09-457-037B-30	Sequence 30, Appl
C 7	49.8	2.1	1501	4	US-09-733-151-30	Sequence 30, Appl
C 8	49	2.1	694	4	US-09-457-037B-39	Sequence 39, Appl
C 9	49	2.1	694	4	US-09-733-151-39	Sequence 39, Appl
C 10	49	2.1	1279	4	US-09-457-037B-40	Sequence 40, Appl
C 11	49	2.1	1279	4	US-09-733-151-40	Sequence 40, Appl
12	48.4	2.1	6317	4	US-10-204-708-11	Sequence 11, Appl
13	48.4	2.1	112132	4	US-09-741-150-3	Sequence 3, Appl
14	48.4	2.1	112132	4	US-10-160-187-3	Sequence 2, Appl
15	47.8	2.0	10467	4	US-10-204-708-2	Sequence 2, Appl
C 16	46	2.0	3095	6	5231168-1	Patent No. 5231168
C 17	45.6	1.9	580073	4	US-08-545-528D-1	Sequence 1, Appl
18	45	1.9	1889	1	US-07-991-867B-41	Sequence 41, Appl
19	45	1.9	1889	2	US-08-544-332-41	Sequence 41, Appl
20	45	1.9	1889	4	US-09-370-861A-41	Sequence 41, Appl
21	45	1.9	1947	4	US-09-370-861A-74	Sequence 74, Appl
22	45	1.9	8457	1	US-07-991-867B-1	Sequence 1, Appl
23	45	1.9	8457	2	US-08-544-332-1	Sequence 1, Appl
24	45	1.9	8457	4	US-09-370-861A-1	Sequence 1, Appl
25	45	1.9	640681	4	US-09-790-988-1	Sequence 1, Appl
C 26	44.8	1.9	4336	4	US-08-956-171E-546	Sequence 546, App
27	44.4	1.9	19124	2	US-08-487-826B-13	Sequence 13, Appl

C 28	44.2	1.9	19124	2	US-08-487-826B-13	Sequence 13, Appl
C 29	44	1.9	832	4	US-09-621-976-2813	Sequence 2813, Ap
C 30	44	1.9	6866	4	US-10-204-708-20	Sequence 20, Appl
C 31	44	1.9	20674	4	US-09-641-638-651	Sequence 651, Appl
C 32	43.8	1.9	1077	4	US-09-457-037B-24	Sequence 24, Appl
C 33	43.8	1.9	1077	4	US-09-733-151-24	Sequence 24, Appl
C 34	43.6	1.9	4140	3	US-08-894-731-2	Sequence 2, Appl
35	43.4	1.8	6113	4	US-10-204-708-14	Sequence 14, Appl
36	43.2	1.8	5841	4	US-10-204-708-89	Sequence 89, Appl
37	43	1.8	8801	4	US-10-204-708-82	Sequence 62, Appl
38	43	1.8	20674	4	US-09-641-638-651	Sequence 651, Appl
39	42.2	1.8	658	3	US-08-998-416-193	Sequence 193, App
40	42.2	1.8	676	3	US-08-998-416-1014	Sequence 1014, Ap
41	42.2	1.8	677	3	US-08-998-416-920	Sequence 920, App
42	42.2	1.8	832	4	US-09-621-976-2813	Sequence 2813, Ap
C 43	42.2	1.8	1274	4	US-09-598-401C-34	Sequence 34, Appl
44	42	1.8	5562	4	US-10-204-708-63	Sequence 63, Appl
45	42	1.8	5666	4	US-10-204-708-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1
US-09-702-705-1269
; Sequence 1269, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Pan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1269
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(675)
; OTHER INFORMATION: n = A,T,C or G
US-09-702-705-1269

Query Match	24.8%;	Score 581.4;	DB 4;	Length 675;
Best Local Similarity	97.3%;	Pred. No. 8.5e-136;		
Matches	62;	Conservative	0;	Mismatches 13; Indels 4; Gaps 3;
Qy	1673	CTGAAAAGAGTGTCTCTCAATATCTCACTAAGTCTCTCACTCAAGCAGATTCT	1732	
Db	1	CTGAAAAGAGTGTCTCTCAATATCTCTCACTAAGTCTCTCACTCAAGCAGATTCT	60	
Qy	1733	TCACCTCGGACATGTGATCATGAACTTAGTAGAGGGATGTGTATTTTATACAAAT	1792	
Db	61	TCACCTCGGACATGTGATCATGAACTTAGTAGAGGGATGTGTATTTTATACAAAT	120	
Qy	1793	TTAATCAATGTCTTACATTTGATTAATCTTAAAGAGCAAACTGATTTTATTTCTGC	1852	
Db	121	TTAATCAATGTCTTACATTTGATTAATCTTAAAGAGCAAACTGATTTTATTTCTGC	180	
Qy	1853	ATCCACATTTCCATCATATTAGAACTTATGATATTTTATCTATGAAGATATAAATGGTGCAG	1912	
Db	181	ATCCACATTTCCATCATATTAGAACTTATGATATTTTATCTATGAAGATATAAATGGTGCAG	240	


```
QY 1673 CTGAAAAGAGTGATCCTCAATATCCTAACTAACTGGTCCCTCAACTCAAGCAGAGTTTCT 1732
Db 1 CTGAAAAGAGTGATCCTCAATATCCTAACTAACTGGTCCCTCAACTCAAGCAGAGTTTCT 60
QY 1733 TCACCTCTGGCAGCTGTGATCATGAAACTTAGTAGAGGGGATTTGTGTATTTTATACAAT 1792
Db 61 TCACCTCTGGCAGCTGTGATCATGAAACTTAGTAGAGGGGATTTGTGTATTTTATACAAT 120
QY 1793 TTAATACANTGCTTACATTTGATTAATTTCTTAAAGAGCAAACTGCAATTTTCTTGC 1852
Db 121 TTAATACANTGCTTACATTTGATTAATTTCTTAAAGAGCAAACTGCAATTTTCTTGC 180
QY 1853 ATCCCAATTTCCCAATCATATTAAGAACTTAAGATATTTTATGAAGATATAAATGGTGCAG 1912
Db 181 ATCCCAATTTCCCAATCATATTAAGAACTTAAGATATTTTATGAAGATATAAATGGTGCAG 240
QY 1913 AGAGACTTTTCATCTGTGGATTCGGTTGTTTCTTAGGGTTCCTAGCACTGATGCCCTGCACA 1972
Db 241 AGAGACTTTTCATCTGTGGATTCGGTTGTTTCTTAGGGTTCCTAGCACTGATGCCCTGCACA 300
QY 1973 AGCATGTGATATGTGAAATAAATGGATTCCTTATAGCTAAATGAGTTCCCTCTGGGGA 2032
Db 301 AGCATGTGATATGTGAAATAAATGGATTCCTTATAGCTAAATGAGTTCCCTCTGGGGA 360
QY 2033 GAGTTCGTGATCTGCAATCAAAATCCAGATGGTGTATAGGGCTATTTGTGTAAGTAAG 2092
Db 361 GAGTTCGTGATCTGCAATCAAAATCCAGATGGTGTATAGGGCTATTTGTGTAAGTAAG 420
QY 2093 TGTGAAGTGTATGAAGTAAGTGTGTTTTCATCTTATGGAATCTTGTATGTCATG 2152
Db 421 TGTGAAGTGTATGAAGTAAGTGTGTTTTCATCTTATGGAATCTTGTATGTCATG 480
QY 2153 TGTCTTTGATGGAATAAATTTTGTGCAATATGATGTCATTCACCTTGTGCAATTTG 2212
Db 481 TGTCTTTGATGGAATAAATTTTGTGCAATATGATGTCATTCACCTTGTGCAATTTG 540
QY 2213 -AATTTTGGTGTATTTATATGATGATATA- CCTGTGACGCTTCTAGTTCCTCAACCAATT 2270
Db 541 AATTTTGGTGTATTTATATGATGATATA- CCTGTGACGCTTCTAGTTCCTCAACCAATT 600
QY 2271 TTATAACCAATTTTGTACATATTTTACTTGAATAATTTT 2309
Db 601 T--ATACCAATTTTGNACATATTTTACTTGNAAATATTT 637
```

RESULT 4

```
US-09-614-124B-1269
; Sequence 1269, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1269
; LENGTH: 665
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1..665)
; OTHER INFORMATION: n = A,T,C or G
```

US-09-614-124B-1269

```
Query Match 23.5%; Score 551.4; DB 4; Length 665;
Best Local Similarity 95.8%; Pred. No. 2.6e-128;
Matches 612; Conservative 0; Mismatches 13; Indels 14; Gaps 4;
QY 1673 CTGAAAAGAGTGATCCTCAATATCCTAACTAACTGGTCCCTCAACTCAAGCAGAGTTTCT 1732
Db 1 CTGAAAAGAGTGATCCTCAATATCCTAACTAACTGGTCCCTCAACTCAAGCAGAGTTTCT 60
QY 1733 TCACCTCTGGCAGCTGTGATCATGAAACTTAGTAGAGGGGATTTGTGTATTTTATACAAT 1792
Db 61 TCACCTCTGGCAGCTGTGATCATGAAACTTAGTAGAGGGGATTTGTGTATTTTATACA-- 117
QY 1793 TTAATCAATGCTTACATTTGATTAATTTCTTAAAGAGCAAACTGCAATTTTCTTGC 1852
Db 118 -----AATGCTTACATTTGATTAATTTCTTAAAGAGCAAACTGCAATTTTCTTGC 170
QY 1853 ATCCCAATTTCCCAATCATATTAAGAACTTAAGATATTTTATGAAGATATAAATGGTGCAG 1912
Db 171 ATCCCAATTTCCCAATCATATTAAGAACTTAAGATATTTTATGAAGATATAAATGGTGCAG 230
QY 1913 AGAGACTTTTCATCTGTGGATTCGGTTGTTTCTTAGGGTTCCTAGCACTGATGCCCTGCACA 1972
Db 231 AGAGACTTTTCATCTGTGGATTCGGTTGTTTCTTAGGGTTCCTAGCACTGATGCCCTGCACA 290
QY 1973 AGCATGTGATATGTGAAATAAATGGATTCCTTATAGCTAAATGAGTTCCCTCTGGGGA 2032
Db 291 AGCATGTGATATGTGAAATAAATGGATTCCTTATAGCTAAATGAGTTCCCTCTGGGGA 350
QY 2033 GAGTTCGTGATCTGCAATCAAAATCCAGATGGTGTATAGGGCTATTTGTGTAAGTAAG 2092
Db 351 GAGTTCGTGATCTGCAATCAAAATCCAGATGGTGTATAGGGCTATTTGTGTAAGTAAG 410
QY 2093 TGTGAAGTGTATGAAGTAAGTGTGTTTTCATCTTATGGAATCTTGTATGTCATG 2152
Db 411 TGTGAAGTGTATGAAGTAAGTGTGTTTTCATCTTATGGAATCTTGTATGTCATG 470
QY 2153 TGTCTTTGATGGAATAAATTTTGTGCAATATGATGTCATTCACCTTGTGCAATTTG 2212
Db 471 TGTCTTTGATGGAATAAATTTTGTGCAATATGATGTCATTCACCTTGTGCAATTTG 530
QY 2213 -AATTTTGGTGTATTTATATGATGATATA- CCTGTGACGCTTCTAGTTCCTCAACCAATT 2270
Db 531 AATTTTGGTGTATTTATATGATGATATA- CCTGTGACGCTTCTAGTTCCTCAACCAATT 590
QY 2271 TTATAACCAATTTTGTACATATTTTACTTGAATAATTTT 2309
Db 591 T--ATACCAATTTTGNACATATTTTACTTGNAAATATTT 627
```

RESULT 5

```
US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
; OTHER INFORMATION: n = A,T,C or G
```

```

Query Match      2.3%; Score 54.6; DB 4; Length 640681;
Best Local Similarity 45.7%; Pred. No. 0.004;
Matches 239; Conservative 0; Mismatches 274; Indels 10; Gaps 1;

QY 475 TAACTTTCTTTATGCTCTGAAAGAGAAAGCAATCAATCACTATGATTAATGACAA 534
Db 574858 TATATATATATTTACTATATAGAGAAATAGTAAAAATAACCTTTAATATCTTAATATAA 574799

QY 535 ACACAGAGATTTCCACAAATTTTAACTGATCTCTCTTTGGAGAAATTTGTTA 594
Db 574798 AAGAGAAATCTAGCTCAAGTTTCTGATTTTGTATATATCTTTTAAAGATATCT 574739

QY 595 CCTAATAGTTTTCCTTATGATGTTTATTACTACTGTTTAAATCAAAATTTCTATAAAT 654
Db 574738 AATATTTTTTTTATAGTAAAGTAAAGAAATTTGTTTTTAAATAAAAATAAGAAAAAT 574679

QY 655 TCCTACTTAAGTCTTAAGACTGGTT-----CTTCTTTGATGTTTATTCATGTT 704
Db 574678 TATAGTAACTGATAGATATATTTTTTAAATAATCAATTTTTTATCTTTTATGTTA 574619

QY 705 CAGAAAGGAAACACACTTTTACTCTTTTAGGACAAATCCCTAGAAATCTATAGTAGTATCAG 764
Db 574618 CTAATAAGAGAGGATTAATTTTAAATGAAAAATAATACAAATTTTTTAAATAATATACAT 574559

QY 765 GATATATTTTGGCTTTAAATATATTTTGGTTATTTTGAATACAGACATTTGGCTCCAAAT 824
Db 574558 CACTCGTTAATTTAAATAATTTTATTAATGTTTTTAAATATAGAAATTTACAAATATAT 574499

QY 825 TTCATCTTTGCACAAATAGTATGACTTTTCTACTAGAACTTCTCAACATTTTGGGAACTTTGC 884
Db 574498 TTTACATTTTAAATGATTTTAAATTTTAAATTTTAAATAATTAATTAATTTTCTTATTTCA 574439

QY 885 AATATGAGCATCATATGCTTAAAGCTGTATCATTTTAAATGCTATGATGATACATTTT 944
Db 574438 ACGTGTCTTATGCTCCGACCAAGATGCTGAGCAITGGTATTATTAAAGAAATATC 574379

QY 945 CTCCTATGCAACAGCTGAAACAAAGCTAGTTGTTTTTACT 987
Db 574378 TTAGTATAAATAATATGCGAGAAACGAATATTTTCTAGT 574336

RESULT 6
US-09-457-037B-30/c
; Sequence 30, Application US/09457037B
; Patent No. 6506963
; GENERAL INFORMATION:
; APPLICANT: De Both, et al.
; TITLE OF INVENTION: HYBRID WINTER OILSEED RAPE AND METHODS FOR PRODUCING SAME
; FILE REFERENCE: 514412-2019
; CURRENT APPLICATION NUMBER: US/09/457,037B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/457,037
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 30
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1501)
; OTHER INFORMATION: "n" stands for any nucleotide a,c,g or t
; NAME/KEY: misc feature
; LOCATION: (1)..(16)
; OTHER INFORMATION: p-GEM-T vector
; NAME/KEY: misc feature
; LOCATION: (1458)..(1501)
; OTHER INFORMATION: p-GEM-T vector
; NAME/KEY: misc feature
; LOCATION: (1)..(1501)
; OTHER INFORMATION: sequence comprising the 3' flanking region of RF-BN1

```

```

US-09-457-037B-30
Query Match      2.1%; Score 49.8; DB 4; Length 1501;
Best Local Similarity 46.3%; Pred. No. 0.0057;
Matches 156; Conservative 2; Mismatches 179; Indels 0; Gaps 0;

QY 1081 TTATGGGTCTTCTATTTCTTGACAGACAGATTTGACTCAATATTTGTAGAGCTTGGTA 1140
Db 883 TGAATTTGCACTTCTGGTTAAAGTAACTAGGTTTGTGATCATTTTTTGTATATTTAAATA 824

QY 1141 GAATGGATTACATGTTAGTATGATGCTAGTAAATGGTTTTTGTATTGACTCAGAAAT 1200
Db 823 AACTATAGAGATTTCTACTTTCTAAATAAAGTAAATATTTTGTATCCACATTTCTTCT 764

QY 1201 TCATCTCAGGATGAATCTTTTATGCTCTTTTATTTAGTAAAGCATATCTGAATTTTACTTTATA 1260
Db 763 TTTATATGCTATGATGATGATTTTATTTATATATTTATGAGAATTTTGATAAAATTTTATA 704

QY 1261 AAGATGGTTTTAGAAAGCTTTGCTAAATAATTTGGCTAGGAATGGTAATTTCAATTTTCA 1320
Db 703 ATATTTTCCATATATAGTTTACTATAGTTTAAATTTCTAGATTTACTTTTTTGACA 644

QY 1321 GTTCCCAAGGGGTAGAAAAATAATATATGTTGTTTATGTTTATGTTTAAACATATTTATAG 1380
Db 643 GCTACCAATATTTCAATCAAAAAATTTCTATCTTGTATTTATATATTAACAATTTTACTACT 584

QY 1381 GTACTATCTATGATGATTTTAAATATTTTTCATTT 1417
Db 583 TTGTTTATGATTTAAATAATTTTGTATTTTATAAAATTTT 547

RESULT 7
US-09-733-151-30/c
; Sequence 30, Application US/09733151
; Patent No. 6563826
; GENERAL INFORMATION:
; APPLICANT: De Both, et al.
; TITLE OF INVENTION: HYBRID WINTER OILSEED RAPE AND METHODS FOR PRODUCING SAME
; FILE REFERENCE: 514412-2019
; CURRENT APPLICATION NUMBER: US/09/733,151
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/457,037
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 30
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1501)
; OTHER INFORMATION: "n" stands for any nucleotide
; OTHER INFORMATION: "g" stands for either nucleotide g or c
; OTHER INFORMATION: "w" stands for nucleotide a or t/u
; OTHER INFORMATION: sequence comprising the 3' flanking region of RF-BN1.
US-09-733-151-30
Query Match      2.1%; Score 49.8; DB 4; Length 1501;
Best Local Similarity 46.3%; Pred. No. 0.0057;
Matches 156; Conservative 2; Mismatches 179; Indels 0; Gaps 0;

QY 1081 TTATGGGTCTTCTATTTCTTGACAGACAGGATTTGACTCAATATTTGTAGAGCTTGGTA 1140
Db 883 TGAATTTGCACTTCTGGTTAAAGTAACTAGGTTTGTGATCATTTTTTGTATATTTAAATA 824

QY 1141 GAATGGATTACATGTTAGTATGATGCTAGTAAATGGTTTTTGTATTGACTCAGAAAT 1200
Db 823 AACTATAGAGATTTCTACTTTCTAAATAAAGTAAATATTTTGTATCCACATTTCTTCT 764

QY 1201 TCATCTCAGGATGAATCTTTTATGCTCTTTTATTTAGTAAAGCATATCTGAATTTTACTTTATA 1260
Db 763 TTTATATGCTATGATGATTTTATTTATATATTTATGAGAATTTTGATAAAATTTTATA 704

```

QY 1261 AAGATGGTTTGAAGAAGCTTGTCTAAATAATTTGGCCCTAGGAATGGTAACCTTCATTTTCA 1320
DB 703 ATATTTTCCATATATAGTTTACTATAGTTAAATTTAAATCTTAAGATTTACTTTTITGACA 644
QY 1321 GTTGCAGGGGTAGAAAAATAATATGTGTGTGTATGTTTATGTTTAAACATATATTAG 1380
DB 643 GCTACCAATATTTCAATCAAAAATTTGTATCTTGTATTTATATACAAATTCATCTTT 584
QY 1381 GTACTATCTATGAATGATTTAAATATTTTTCATATT 1417
DB 583 TTGTTTCATGATTTKAAAAATTTTAGTTTATAAAATTTT 547

RESULT 8

US-09-457-037B-39/c
; Sequence 39, Application US/09457037B
; Patent No. 6506963
; GENERAL INFORMATION:
; APPLICANT: De Both, et al.
; TITLE OF INVENTION: HYBRID WINTER OILSEED RAPE AND METHODS FOR PRODUCING SAME
; FILE REFERENCE: 514412-2019
; CURRENT APPLICATION NUMBER: US/09/457,037B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/457,037
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39
; LENGTH: 694
; TYPE: DNA
; ORGANISM: Brassica napus
; NAME/KEY: misc feature
; LOCATION: (1)..(694)
; OTHER INFORMATION: sequence comprising the 5' flanking region of RF-BN1 in WOSR
US-09-457-037B-39

Query Match 2.1%; Score 49; DB 4; Length 694;
Best Local Similarity 46.6%; Pred. No. 0.0067;
Matches 157; Conservative 0; Mismatches 180; Indels 0; Gaps 0;
QY 1081 TTATTCGGTCTTCATTTCTGACAGACAGGATTTGACTCAATATTTGAGAGCTTGCCTA 1140
DB 462 TGAATGGACTTCTGGTTAAAGTAACTAGTTTGTGATCAATTTTGCTATATTTAAATA 403
QY 1141 GAATGATTACATGGTAGTGATGCACTGGTAGAAATGTTTTAGTTATTCGACTCAGAAAT 1200
DB 402 AAACATATAAGATTTCTACTCTTAAATAAACTGTAATAATTTTGATCCACATTTGCTTCTT 343
QY 1201 TCATCTCAGGATGAATCTTTTATGCTTTTATTTGTAAGCATATCTGAAATTTACTTTATA 1260
DB 342 TTTATATGCTATGAGTTAGATTTATTTTATATATTTATTTATGAGNAATTTGATAAAATTTATA 283
QY 1261 AAGATGGTTTGAAGAAGCTTGTCTAAATAATTTGGCCCTAGGAATGGTAACCTTCATTTTCA 1320
DB 282 ATATTTTCAATATATAGTTTACTATAGTTAAATTTAAATCTTAAGATTTACTTTTITGACA 223
QY 1321 GTTGCAGGGGTAGAAAAATAATATGTGTGTGTATGTTTATGTTTAAACATATATTAG 1380
DB 222 GCTACCAATATTTCAATCAAAAATTTGTATCTTCTGATTTATATATACAAATTCATCTTT 163
QY 1381 GTACTATCTATGAATGATTTAAATATTTTTCATATT 1417
DB 162 TTGTTTCATGATTTAAAAATTTTAGTTTATAAAATTTT 126

RESULT 9

US-09-733-151-39/c
; Sequence 39, Application US/09733151
; Patent No. 6563026
; GENERAL INFORMATION:
; APPLICANT: De Both, et al.

; TITLE OF INVENTION: HYBRID WINTER OILSEED RAPE AND METHODS FOR PRODUCING SAME
; FILE REFERENCE: 514412-2019
; CURRENT APPLICATION NUMBER: US/09/733,151
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/457,037
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39
; LENGTH: 694
; TYPE: DNA
; ORGANISM: Brassica napus
; NAME/KEY: misc feature
; LOCATION: (1)..(694)
; OTHER INFORMATION: sequence comprising the 5' flanking region of RF-BN1 in WOSR
US-09-733-151-39

Query Match 2.1%; Score 49; DB 4; Length 694;
Best Local Similarity 46.6%; Pred. No. 0.0067;
Matches 157; Conservative 0; Mismatches 180; Indels 0; Gaps 0;
QY 1081 TTATTCGGTCTTCATTTCTGACAGACAGGATTTGACTCAATATTTGAGAGCTTGCCTA 1140
DB 462 TGAATGGACTTCTGGTTAAAGTAACTAGTTTGTGATCAATTTTGCTATATTTAAATA 403
QY 1141 GAATGATTACATGGTAGTGATGCACTGGTAGAAATGTTTTAGTTATTCGACTCAGAAAT 1200
DB 402 AAACATATAAGATTTCTACTCTTAAATAAACTGTAATAATTTTGATCCACATTTGCTTCTT 343
QY 1201 TCATCTCAGGATGAATCTTTTATGCTTTTATTTGTAAGCATATCTGAAATTTACTTTATA 1260
DB 342 TTTATATGCTATGAGTTAGATTTATTTTATATATTTATTTATGAGNAATTTGATAAAATTTATA 283
QY 1261 AAGATGGTTTGAAGAAGCTTGTCTAAATAATTTGGCCCTAGGAATGGTAACCTTCATTTTCA 1320
DB 282 ATATTTTCAATATATAGTTTACTATAGTTAAATTTAAATCTTAAGATTTACTTTTITGACA 223
QY 1321 GTTGCAGGGGTAGAAAAATAATATGTGTGTGTATGTTTATGTTTAAACATATATTAG 1380
DB 222 GCTACCAATATTTCAATCAAAAATTTGTATCTTCTGATTTATATATACAAATTCATCTTT 163
QY 1381 GTACTATCTATGAATGATTTAAATATTTTTCATATT 1417
DB 162 TTGTTTCATGATTTAAAAATTTTAGTTTATAAAATTTT 126

RESULT 10

US-09-457-037B-40/c
; Sequence 40, Application US/09457037B
; Patent No. 6506963
; GENERAL INFORMATION:
; APPLICANT: De Both, et al.
; TITLE OF INVENTION: HYBRID WINTER OILSEED RAPE AND METHODS FOR PRODUCING SAME
; FILE REFERENCE: 514412-2019
; CURRENT APPLICATION NUMBER: US/09/457,037B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/457,037
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 40
; LENGTH: 1279
; TYPE: DNA
; ORGANISM: Brassica napus
; NAME/KEY: misc feature
; LOCATION: (1)..(1279)
; OTHER INFORMATION: sequence comprising the 3' flanking region of RF-BN1 in WOSR
US-09-457-037B-40
Query Match 2.1%; Score 49; DB 4; Length 1279;
Best Local Similarity 46.6%; Pred. No. 0.0085;


```
Db 4210 AATAAGATATAGTTTTATAGATTATATATGTTAGTGCAGAGTTTT 4258
RESULT 13
US-09-741-150-3
; Sequence 3, Application US/09741150
; Patent No. 6436689
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: US$ THEREOF
; FILE REFERENCE: CL000968
; CURRENT APPLICATION NUMBER: US/09/741,150
; CURRENT FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 112132
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(112132)
; OTHER INFORMATION: n = A,T,C or G
US-09-741-150-3
Query Match 2.1%; Score 48.4; DB 4; Length 112132;
Best Local Similarity 55.3%; Pred. No. 0.07;
Matches 94; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
QY 720 ACTTTACTCTTTAGGACAAATCCTAGATCTATAGTATCAGGATATATTTGCTTT 779
Db 26831 ACTTCTACCTAGAGCTGCTGCTAAATCAATAGTATTAATTAATATAGTT 26890
QY 780 AAAATATATTTGGTATTTTGAATACAGACATTTGGCTCCAAATTTTCATCTTTGCACAA 839
Db 26891 ACTGCTTAGCTAGGTTTTTTGGGGGACTAGCTTGGGAACCAATACCACTCAGGCCAT 26950
QY 840 TAGTATGACTTTTCACATAGACATCTCTCAACATTTGGGAACCTTTGCAATA 889
Db 26951 TTTTTCCTTTATGAATATATCCTTAGCAAAATTTAAATTAATTAATAA 27000
RESULT 15
US-10-204-708-2
; Sequence 2, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/BP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 2
; LENGTH: 10467
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-2
Query Match 2.0%; Score 47.8; DB 4; Length 10467;
Best Local Similarity 45.6%; Pred. No. 0.039;
Matches 169; Conservative 0; Mismatches 202; Indels 0; Gaps 0;
QY 289 TTTCGACTGTCTTATTTCAAAAGATTTCTCAATTAATTTTCCCTAGAAATTTCTCCCTCA 348
Db 7895 TTATGCTTTTATATATAAGTTTTTTTATCGTTAGTATTTATAGTATATAT 7954
QY 349 TTCAAAGTCAACATACCTTTGAAGAAATGAACAGATTTCCCATGATGATGCTCA 408
Db 7955 ATTTTTTTGATTAATAATATATTTGAATTTTAAAAAATTTATATTTTAAATTTT 8014
QY 409 TACTCGACTAGAACGATCTATGTTAAATGACGTGTATATGAATTTTCAAGTACTAC 468
Db 8015 TGTTATGTTTAAATGTTGTTTTTGTAAATAGTATTTTGAATTAATTAATTAATTG 8074
QY 469 CCCAATTAACCTTTCTTATTCCTCTGAAAGAAAGCAATGTAATCACTATGATPAT 528
Db 8075 GTTTGATAAATTTTTTTTTTTGTTGGAGTTTATAATTTTATGCGAATATCGAAATAT 8134
QY 529 GCACAAACACCAAGATTTCTCCACAAATTTTAAATTAATCTGATCTCTTCTCGAGAAA 588
Db 8135 TTGGATATATTTATATATTTTCTGTTTATTTTATTTTATTTTATTTTACGTTTTTTT 8194
QY 589 TTGTTACCTAATAGTTTTTTCCTTATGAATGTTTATTACTACTGTTATAAATCAAAATTTCTA 648
```

Db 8195 ATTTTAAATTTTTCGANTGATTTTATTTTATTTTAAATTTTATAGTT 8254
Qy 649 TAAATTCCTA 659
Db 8255 TCAAAGTTGTA 8265

Search completed: April 26, 2004, 02:11:20
Job time : 120.382 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2004, 23:47:42 ; Search time 623.343 Seconds
(without alignments)
15988.422 Million cell updates/sec

Title: US-10-051-835-25

Perfect score: 2346

Sequence: 1 gccagaggggaaaaaagag.....aaacatttgatgttaccat 2346

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 337863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: N Geneseq 29Jan04:*

- 1: Geneseqm1980as:*
- 2: Geneseqm1990as:*
- 3: Geneseqm2000as:*
- 4: Geneseqm2001as:*
- 5: Geneseqm2001bs:*
- 6: Geneseqm2002as:*
- 7: Geneseqm2003as:*
- 8: Geneseqm2003bs:*
- 9: Geneseqm2003cs:*
- 10: Geneseqm2004as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2346	100.0	2346	8	ACH04269 Human CDN
2	2334	99.5	2350	9	ADE25669 Human CDN
3	2304.4	98.2	2321	6	AAS94792 Human DNA
4	1599.4	68.2	1626	3	AAS98844 Human dif
5	1599.4	68.2	1626	3	AAS98125 Human dif
6	1599.4	68.2	1626	6	AAD29955 Human pha
7	603.2	25.3	665	4	AAS34083 Human col
8	593	25.3	603	6	ABZ08185 Human leu
9	584.8	24.9	608	4	AAL22443 Human bre
10	582.8	24.8	586	6	ABK27756 Human col
11	581.4	24.8	675	6	ABK39231 DNA encod
12	581.4	24.8	675	7	ACAL1560 Human lun
13	581.4	24.8	675	7	ACA02746 Lung can
14	568	24.2	586	6	ABU37834 Human col
15	546.2	23.3	783	4	AAU13574 Human bre
16	497.4	21.2	499	6	ABV87830 Human col
17	483	20.6	483	6	ABV87118 Human col
18	468	19.9	482	4	AAS36474 Human col
19	435	18.5	435	6	ABV87338 Human col
20	401.8	17.1	2000	7	ABU17369 Human SLC
21	323	13.8	323	6	ABU65495 Lung can
22	323	13.8	323	6	ABK64193 Human ben
23	323	13.8	323	6	ABN93962 Gene #460

C	24	322.2	13.7	324	6	ABL38369	Human col
C	25	321.2	13.7	324	6	ABL37940	Human col
C	26	321	13.7	321	4	AAD23574	Human lun
C	27	321	13.7	321	9	ADD66862	Human lun
C	28	321	13.7	321	9	ADBE8116	Human lun
C	29	319.6	13.6	324	6	ABL38210	Human col
C	30	273.2	11.6	277	6	ABL38570	Human col
C	31	258.4	11.0	284	2	AAT25840	Human gen
C	32	243.6	10.4	247	6	ABL37563	Human col
C	33	242.6	10.3	247	6	ABL38537	Human col
C	34	242.6	10.3	247	6	ABL38589	Human col
C	35	242.4	10.3	244	6	ABK27670	Human col
C	36	229.6	9.8	248	6	ABL37833	Human col
C	37	64	2.7	309	6	ABL38438	Human col
C	38	61	2.6	8056	7	ABZ10246	Haematopo
C	39	59.2	2.5	700	4	AAS93026	Human inf
C	40	58.8	2.5	5982	6	ABL34049	Human imm
C	41	57.4	2.4	8056	7	ABZ10246	Haematopo
C	42	57.4	2.4	18218	6	ABL33949	Human imm
C	43	57.2	2.4	13584	6	ABL32615	Human imm
C	44	57.2	2.4	17538	6	ABL33157	Human imm

ALIGNMENTS

RESULT 1
ACH04269

ID ACH04269 standard; cDNA; 2346 BP.

AC ACH04269;

DT 26-SEP-2003 (first entry)

DE Human cDNA differentially expressed in lung cancer #474.

XX Gene therapy; emphysema; ss; gene; chronic obstructive pulmonary disease;
XX respiratory disorder; lung cancer; asthma; human.

OS Homo sapiens.

XX US2003065157-A1.

XX 03-APR-2003.

DF 04-APR-2002; 2002US-00116802.

PR 04-APR-2001; 2001US-0281593P.

XX (LASE/) LASEK A W.

XX Lasek AW;

XX WPI; 2003-540803/51.

XX New combination comprising cDNAs that are differentially expressed in
PT respiratory disorders, useful for diagnosing or treating respiratory
PT disorders e.g., lung cancer, chronic obstructive pulmonary disease,
PT emphysema or asthma.

PS Claim 1; Page; 39pp; English.

XX The invention relates to a combination comprising cDNAs or their
CC complements that are differentially expressed in respiratory disorder.
CC The combination is useful for preparing a composition for diagnosing or
CC treating respiratory disorders e.g., lung cancer, chronic obstructive
CC pulmonary disease, emphysema or asthma. The present sequence represents
CC human cDNA differentially expressed during lung cancer

SQ Sequence 2346 BP; 712 A; 361 C; 428 G; 845 T; 0 U; 0 Other;

Query Match 100.0%; Score 2346; DB 8; Length 2346;

	Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches	0;	Indels	0;	Gaps	0;
	Matches	2346;	Conservative	0;					
QY	1	GC	CAGAGGGG	AAAAAAGAGTAATGCACAGGTATCTCTTTTGCAGTGGGACGTATTTT	60				
Db	1	GC	CAGAGGGG	AAAAAAGAGTAATGCACAGGTATCTCTTTTGCAGTGGGACGTATTTT	60				
QY	61	GAGT	ACTCTT	GTGTGACAGGGTATTATTACAGCATCTTGTGGGAAAAACCTATTAGGCGCTT	120				
Db	61	GAGT	ACTCTT	GTGTGACAGGGTATTATTACAGCATCTTGTGGGAAAAACCTATTAGGCGCTT	120				
QY	121	TG	CATGTTT	TAAGCTGTATTAATTTTGTGGTGTGAGTGGTCTGCACCTAAATCTGTATTAT	180				
Db	121	TG	CATGTTT	TAAGCTGTATTAATTTTGTGGTGTGAGTGGTCTGCACCTAAATCTGTATTAT	180				
QY	181	AAA	ATT	TAGACATCAAAATTTTCTCTACTAACTAACTTTATTAGATGCATACTTTGGGAAGCAC	240				
Db	181	AAA	ATT	TAGACATCAAAATTTTCTCTACTAACTAACTTTATTAGATGCATACTTTGGGAAGCAC	240				
QY	241	AGT	CATAT	TCACACTGGGAGCAATGCAATGTGTGTACCTGGTCTTACCTAGTCTTGAATCTGTCT	300				
Db	241	AGT	CATAT	TCACACTGGGAGCAATGCAATGTGTGTGTACCTGGTCTTACCTAGTCTTGAATCTGTCT	300				
QY	301	TAT	TT	CAAAAAGATTTCTGAAATTAATTTTCCCTAGAAATTTCTCTTCAATCCAAAGTACA	360				
Db	301	TAT	TT	CAAAAAGATTTCTGAAATTAATTTTCCCTAGAAATTTCTCTTCAATCCAAAGTACA	360				
QY	361	AAC	ATAT	TTGAAGATGAACAGATGTGTCCCATGATGTATGCTCATACTCGACTAGA	420				
Db	361	AAC	ATAT	TTGAAGATGAACAGATGTGTCCCATGATGTATGCTCATACTCGACTAGA	420				
QY	421	AAC	GATCTAT	TGTTAAATGACTGTGATATGAATTTATTTCAAGTACTACCCAAATAACTTT	480				
Db	421	AAC	GATCTAT	TGTTAAATGACTGTGATATGAATTTATTTCAAGTACTACCCAAATAACTTT	480				
QY	481	TC	TAT	TGCTCTGAAAGAGAAAGCAATGTAATCAGTATGATTATGACAAAACAAC	540				
Db	481	TC	TAT	TGCTCTGAAAGAGAAAGCAATGTAATCAGTATGATTATGACAAAACAAC	540				
QY	541	AGA	TTCT	CTCAACAATTTAAGTAATCTGATCTCTTTTGGAGAAATTTGTTACTTAAT	600				
Db	541	AGA	TTCT	CTCAACAATTTAAGTAATCTGATCTCTTTTGGAGAAATTTGTTACTTAAT	600				
QY	601	AG	TTTT	CTTTATCAAGTATTATCTACTCGGTATTAATCAAAATTTCTATATAATTTCTCTAC	660				
Db	601	AG	TTTT	CTTTATCAAGTATTATCTACTCGGTATTAATCAAAATTTCTATATAATTTCTCTAC	660				
QY	661	TTA	AGTCT	TAAAGACTGGGTCTTCTCTTTGATGTTATTCATGTTTCAGAAAGGAAAACA	720				
Db	661	TTA	AGTCT	TAAAGACTGGGTCTTCTCTTTGATGTTATTCATGTTTCAGAAAGGAAAACA	720				
QY	721	CT	TTT	ACTCTTTTAGGCAATTCCTAGAACTCTATAGTAGTATCAGGATATATTTTGCTTTA	780				
Db	721	CT	TTT	ACTCTTTTAGGCAATTCCTAGAACTCTATAGTAGTATCAGGATATATTTTGCTTTA	780				
QY	781	AA	TAT	ATATTTTGGTATTATTTTGAATACAGACATGGCTCCAAATTTTCATTTTGCACAAT	840				
Db	781	AA	TAT	ATATTTTGGTATTATTTTGAATACAGACATGGCTCCAAATTTTCATTTTGCACAAT	840				
QY	841	AG	TAT	GTCTTCTCACTAGAACCTCTCAACATTTTGGGAACCTTTCGAAATATGAGCATCAT	900				
Db	841	AG	TAT	GTCTTCTCACTAGAACCTCTCAACATTTTGGGAACCTTTCGAAATATGAGCATCAT	900				
QY	901	TG	TG	TAAAGCTGTATCAATTTAAATGCTATGAGATACATTTGTTTCTCCCTATGCGCAACA	960				
Db	901	TG	TG	TAAAGCTGTATCAATTTAAATGCTATGAGATACATTTGTTTCTCCCTATGCGCAACA	960				
QY	961	GG	TGA	CAAAACGTAGTCTTTTTTTTACTGATATAAAATGTTGGCTACCTGTGATTTTATAG	1020				
Db	961	GG	TGA	CAAAACGTAGTCTTTTTTTTACTGATATAAAATGTTGGCTACCTGTGATTTTATAG	1020				
QY	1021	TAT	G	CACATGTGCAGAAAAAGGCAAGACAAATTTGGCTCTTGTACTGATATCTTTGGGAAAC	1080				

Db	1021	TATGACATGTCAGAAAAGCGACAGCAAAATGCGCTCTTGTA	CTGAAATACTTTCCGGCAAAAC	1080	
Qy	1081	TTATTGGGTCCTTCATTTTCTGACAGACAGAGATTGTGACT	CAATATTTGTAGAGCTTGCCTA	1140	
Db	1081	TTATTGGGTCCTTCATTTTCTGACAGACAGAGATTGTGACT	CAATATTTGTAGAGCTTGCCTA	1140	
Qy	1141	GAATGGATTACATGGTAGTGATGCACCTGGTAGAAAATGGT	TTTTAGTTATTGACTCAGAAT	1200	
Db	1141	GAATGGATTACATGGTAGTGATGCACCTGGTAGAAAATGGT	TTTTAGTTATTGACTCAGAAT	1200	
Qy	1201	TCATCTCAGAGTGAATCTTTTATGCTCTTTTATATGTAAG	CATATCTGAATTTTACTTTTAA	1260	
Db	1201	TCATCTCAGAGTGAATCTTTTATGCTCTTTTATATGTAAG	CATATCTGAATTTTACTTTTAA	1260	
Qy	1261	AAGATGGTTTTAGAAAGCTTTGCTCTAAAAAATTTGGCCT	TAGGAATGGTAACTTTTCTCA	1320	
Db	1261	AAGATGGTTTTAGAAAGCTTTGCTCTAAAAAATTTGGCCT	TAGGAATGGTAACTTTTCTCA	1320	
Qy	1321	GTTGCCAAGGGGTAGAAAATTAATATGCTGTGTTTATGTT	TATGTTTAAACATATTTATAG	1380	
Db	1321	GTTGCCAAGGGGTAGAAAATTAATATGCTGTGTTTATGTT	TATGTTTAAACATATTTATAG	1380	
Qy	1381	GTACTATCTATGAATGTATTTAAATATTTTTCATATCTGT	GACAAGCATTTTATATTTTG	1440	
Db	1381	GTACTATCTATGAATGTATTTAAATATTTTTCATATCTGT	GACAAGCATTTTATATTTTG	1440	
Qy	1441	CAACRAGTGGAGTCACATTTAGCCCGACGTGGGAAAGTCT	TGGAACTCTTGGAACTTACCG	1500	
Db	1441	CAACRAGTGGAGTCACATTTAGCCCGACGTGGGAAAGTCT	TGGAACTCTTGGAACTTACCG	1500	
Qy	1501	GATATGCTGGCAGCCATCTCTTTGATGCTGTGCTTAAACT	TGTAAATTTTATAGACAGCTAAA	1560	
Db	1501	GATATGCTGGCAGCCATCTCTTTGATGCTGTGCTTAAACT	TGTAAATTTTATAGACAGCTAAA	1560	
Qy	1561	TCCTTAACCTTGGATCTGGAATGCATTAGTTATGACCTTG	TACCATTTCCCAGAAATTTTCAGG	1620	
Db	1561	TCCTTAACCTTGGATCTGGAATGCATTAGTTATGACCTTG	TACCATTTCCCAGAAATTTTCAGG	1620	
Qy	1621	GGCATCGTGGTTTGGTCTAGTGATTTGAAACAACAAGAA	CACAGAGATCCAGCTGGAATAA	1680	
Db	1621	GGCATCGTGGTTTGGTCTAGTGATTTGAAACAACAAGAA	CACAGAGATCCAGCTGGAATAA	1680	
Qy	1681	GAGTGATCCCTCAATATCCCTAACCTTAACTGGTCCCTCA	ACTCAAGCAGAGATTTCTTCAC	1740	
Db	1681	GAGTGATCCCTCAATATCCCTAACCTTAACTGGTCCCTCA	ACTCAAGCAGAGATTTCTTCAC	1740	
Qy	1741	GCATCTGTGATCATGAACCTTAGTAGAGGGGATTTGTGTG	TATTTTATACAAATTTTATACA	1800	
Db	1741	GCATCTGTGATCATGAACCTTAGTAGAGGGGATTTGTGTG	TATTTTATACAAATTTTATACA	1800	
Qy	1801	ATGTCTTACATTCGATAAAAATCTTTAAAGAGCAAAAAC	TGCAATTTTATTTCTGCAATCCACAT	1860	
Db	1801	ATGTCTTACATTCGATAAAAATCTTTAAAGAGCAAAAAC	TGCAATTTTATTTCTGCAATCCACAT	1860	
Qy	1861	TCCAAATCATTTAGAACTAAGATATTTATCTATGAAGAT	ATAATGGTGCAGAGAGACTTT	1920	
Db	1861	TCCAAATCATTTAGAACTAAGATATTTATCTATGAAGAT	ATAATGGTGCAGAGAGACTTT	1920	
Qy	1921	TCATCTGTGATTCGCTGTGTTTCTTAGGGTTCCCTAGCA	CTGATCCCTGCAACAAGCATGTG	1980	
Db	1921	TCATCTGTGATTCGCTGTGTTTCTTAGGGTTCCCTAGCA	CTGATCCCTGCAACAAGCATGTG	1980	
Qy	1981	ATATGTGAAATTAATAATGGATCTTCTATAGCTAAATG	AGTTCCCTCTGGGGAGAGTTCTG	2040	
Db	1981	ATATGTGAAATTAATAATGGATCTTCTATAGCTAAATG	AGTTCCCTCTGGGGAGAGTTCTG	2040	
Qy	2041	GTA	CTGCAATCACAAATGCCAGATGGTCTTTATATGGGCTAT	TTTGTGTAAAGTAAGTGGTAAGA	2100
Db	2041	GTA	CTGCAATCACAAATGCCAGATGGTCTTTATATGGGCTAT	TTTGTGTAAAGTAAGTGGTAAGA	2100
Qy	2101	TGCTATGAAGTAAGTGTTGTTTTCATCTTATGGAACCT	CTTATGTAAGCATGTGCTTTTG	2160	
Db	2101	TGCTATGAAGTAAGTGTTGTTTTCATCTTATGGAACCT	CTTATGTAAGCATGTGCTTTTG	2160	

QY 2161 TATGGAATAAATTTGCTGCAATATGATGTCATTCACCTTTGCAATGAAATTTGG 2220
 Db 2161 TATGGAATAAATTTGCTGCAATATGATGTCATTCACCTTTGCAATGAAATTTGG 2220
 QY 2221 TTGTATTTATATGATATATACCTGTCACGGCTTCTAGTTGCTTCAACCATTTTATACCAT 2280
 Db 2221 TTGTATTTATATGATATATACCTGTCACGGCTTCTAGTTGCTTCAACCATTTTATACCAT 2280
 QY 2281 TTTTGTACATATTTTACTTGAAATATTTTAAATGGAATTTTAAATACATTTGTATAGT 2340
 Db 2281 TTTTGTACATATTTTACTTGAAATATTTTAAATGGAATTTTAAATACATTTGTATAGT 2340
 QY 2341 TTACAT 2346
 Db 2341 TTACAT 2346

RESULT 2

AD25669
 ID AD25669 standard; cDNA; 2350 BP.

AC AD25669;

XX 29-JAN-2004 (first entry)

DE Human cDNA differentially expressed in foam cells #73.

XX Human; ss; differential expression; foam cell; LPS; lipopolysaccharide;
 KW Cardiovascular disease; atherosclerosis.

XX Homo sapiens.

XX US2003194721-A1.

XX 16-OCT-2003.

XX 18-SEP-2002; 2002US-00247671.

XX 19-SEP-2001; 2001US-0323784P.

XX (INCY-) INCYTE GENOMICS INC.

XX Mikita T, Shiffman D, Porter JG, Kaser MR;

XX WPI; 2003-875398/81.

XX Combination containing several polynucleotide that are differentially
 PT expressed in foam cells and complements of the polynucleotides, useful
 PT for diagnosing cardiovascular disease or atherosclerosis.

XX Claim 1; SEQ ID NO 73; 37pp; English.

XX The invention relates to a combination comprising several polynucleotides
 CC having any one of 127 sequences (S1) such as the sequence of human
 CC calmodulin gene, human mRNA for KIA0930 protein, leukotriene A4
 CC hydrolase, human CG1-142 protein mRNA, human K⁺ channel beat 2 subunit
 CC mRNA, etc., and their complements. The cDNAs are differentially expressed
 CC in LPS (lipopolysaccharide)-treated foam cells. Also included are
 CC obtaining an extended or full length gene from a library of nucleic acid
 CC sequences, an expression vector containing the nucleic acids, a host cell
 CC containing the vector, a purified polypeptide appearing as AD25750 and
 CC AD25751, producing a protein by culturing the host cell, and a
 CC composition comprising a purified antibody that specifically binds to the
 CC proteins. The foam cell-expressed nucleic acids are useful for a high
 CC throughput detection of differential expression of one or more
 CC polynucleotides in a sample. The sample is from a subject with
 CC atherosclerosis and comparison with a standard defines early, mid or late
 CC stages of the disorder. The foam cell-expressed nucleic acids are useful
 CC for high throughput screening of a library of molecules or compounds to
 CC identify a ligand which binds a polynucleotide. The library is chosen
 CC from DNA molecules, peptides, proteins and RNA molecules. The protein is
 CC useful for a high throughput screening of library of molecules or

CC compounds to identify at least one ligand which specifically binds a
 CC protein, for purifying a ligand from a sample for making an antibody. The
 CC foam cell-expressed nucleic acids are useful for diagnosing
 CC cardiovascular disorder. The foam cell-expressed nucleic acids are useful
 CC as elements on a microarray which can be used for detecting related
 CC polynucleotide in a sample, diagnosing cardiovascular disease, expression
 CC atherosclerosis. The present sequence represents a cDNA whose expression
 CC is upregulated in LPS treated foam cells.

XX SQ Sequence 2350 BP; 716 A; 361 C; 428 G; 845 T; 0 U; 0 Other;

Query Match 99.5%; Score 2334; DB 9; Length 2350;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2345; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GCCAGAGGGGAAAAAGAGTAATGACAGGGTATATTACAGCATCTTGTGGAAAACTATTAGGCCCTT 60
 Db 1 GCCAGAGGGGAAAAAGAGTAATGACAGGGTATATTACAGCATCTTGTGGAAAACTATTAGGCCCTT 60
 QY 61 GAGTACCTTG-TGTGACAGGGTATTATTACAGCATCTTGTGGAAAACTATTAGGCCCTT 119
 Db 61 GAGTACCTTG-TGTGACAGGGTATTATTACAGCATCTTGTGGAAAACTATTAGGCCCTT 119
 QY 121 TGCATGTTTAAGCTGTATAAATTTGGTGTGAGTGGTCTGACTTAAATGTGTATTAT 180
 Db 120 TGCATGTTTAAGCTGTATAAATTTGGTGTGAGTGGTCTGACTTAAATGTGTATTAT 179
 QY 181 AAAATTTAGACATCAAAATTTCTACTAACTAATCTTATTAGATGCATACATTGGAAGCAC 240
 Db 180 AAAATTTAGACATCAAAATTTCTACTAACTAATCTTATTAGATGCATACATTGGAAGCAC 239
 QY 241 AGTCATATCACACTGGGAGGCAATGCAATGGTTTACCTGGTCTAGGTTTGAAGTGTCT 300
 Db 240 AGTCATATCACACTGGGAGGCAATGCAATGGTTTACCTGGTCTAGGTTTGAAGTGTCT 299
 QY 301 TATTTCAAAGAAATTTCTGAAATTAATTTTCCCTAGAAATTTCTCTTCTTCAATCCAAAGTACA 360
 Db 300 TATTTCAAAGAAATTTCTGAAATTAATTTTCCCTAGAAATTTCTCTTCTTCAATCCAAAGTACA 359
 QY 361 AACATACCTTTGAGATGAAACAGAGATTGTTCCCAATGATGCTGCTATCTGACTAGAGA 420
 Db 360 AACATACCTTTGAGATGAAACAGAGATTGTTCCCAATGATGCTGCTATCTGACTAGAGA 419
 QY 421 AACGATCTATGTTAAATGACTGTGTATATGAATTTTCAAGTACTACCCCAATAAATCTT 480
 Db 420 AACGATCTATGTTAAATGACTGTGTATATGAATTTTCAAGTACTACCCCAATAAATCTT 479
 QY 481 TCTTATTTGCTGGAAGAGAAAGCAATGTAATCACTGTGATTTATTCGACAAACACC 540
 Db 480 TCTTATTTGCTGGAAGAGAAAGCAATGTAATCACTGTGATTTATTCGACAAACACC 539
 QY 541 AGAATTTCTCCAAATTTTAACTAATCTGATCTCTTCTTGGAGAAAATTTGTACCTAAT 600
 Db 540 AGAATTTCTCCAAATTTTAACTAATCTGATCTCTTCTTGGAGAAAATTTGTACCTAAT 599
 QY 601 AGTTTTCTCTTATGATGCTGTTATCTGCTGATGTAATCAATTTCTATTAATTTCTCTAC 660
 Db 600 AGTTTTCTCTTATGATGCTGTTATCTGCTGATGTAATCAATTTCTATTAATTTCTCTAC 659
 QY 661 TTAAGTCTTTAAGAACTGGGTTTCTTCTTTGATGTTTATTCTGTTTCAGAAAGAAACACA 720
 Db 660 TTAAGTCTTTAAGAACTGGGTTTCTTCTTTGATGTTTATTCTGTTTCAGAAAGAAACACA 719
 QY 721 CTTTACTCTTTTAGGCAATTTCTAGATCTATAGTATCAGGATATATTTTGCCTTTA 780
 Db 720 CTTTACTCTTTTAGGCAATTTCTAGATCTATAGTATCAGGATATATTTTGCCTTTA 779
 QY 781 AAATATATTTTGGTTTATTGTAATACAGACATTTGGCTCCAAATTTTCACTTTTGCACAAAT 840
 Db 780 AAATATATTTTGGTTTATTGTAATACAGACATTTGGCTCCAAATTTTCACTTTTGCACAAAT 839
 QY 841 AGTATGACCTTTTCACTAGACCTTCTCAACATTTGGGACCTTTGCAATATGAGCATCATA 900

Db 840 AGTATGACCTTTTCACAGACCTTCTCAACATTTGGGAACCTTGCAAAATATGAGCATCAT 899
Qy 901 TGTGTTAAGGCTGTATCATTTAATGCTATGAGATACATTTGTTTCTCCCTATGCCAAACA 960
Db 900 TGTGTTAAGGCTGTATCATTTAATGCTATGAGATACATTTGTTTCTCCCTATGCCAAACA 959
Qy 961 GGTGACAAACCTAGTGTGTTTTTACTGATCACTAAATGTTGGCTACCTGTGATTTTATAG 1020
Db 960 GGTGACAAACCTAGTGTGTTTTTACTGATCACTAAATGTTGGCTACCTGTGATTTTATAG 1019
Qy 1021 TATGACATGTGAGAAAAGGCAAGACAAATGGCTCTTGTACTGAATACCTTCGGCAAC 1080
Db 1020 TATGACATGTGAGAAAAGGCAAGACAAATGGCTCTTGTACTGAATACCTTCGGCAAC 1079
Qy 1081 TTATTGGGCTTCATTTCTGACAGACAGAGATTTGACATCAATATTTGAGAGCTTGGCTA 1140
Db 1080 TTATTGGGCTTCATTTCTGACAGACAGAGATTTGACATCAATATTTGAGAGCTTGGCTA 1139
Qy 1141 GAATGATTTACATGTTAGTGTGATGCTGATGCTGAGAAATGGTTTTTACTGATCAGAAAT 1200
Db 1140 GAATGATTTACATGTTAGTGTGATGCTGATGCTGAGAAATGGTTTTTACTGATCAGAAAT 1199
Qy 1201 TCATCTCAGATGAATCTTTTATGCTTTTATGCTAGCATATCTGAAATTTACTTTATA 1260
Db 1200 TCATCTCAGATGAATCTTTTATGCTTTTATGCTAGCATATCTGAAATTTACTTTATA 1259
Qy 1261 AAGATGGTTTTAGAAAGCTTTGCTAAAAAATTTGGCTAGGAATGGTAACCTTCATTTTCA 1320
Db 1260 AAGATGGTTTTAGAAAGCTTTGCTAAAAAATTTGGCTAGGAATGGTAACCTTCATTTTCA 1319
Qy 1321 GTTGCCAAAGGGTAGAAAAATATATATGCTGTGTTGTTATGTTTATGTTAAACATTTATTAG 1380
Db 1320 GTTGCCAAAGGGTAGAAAAATATATATGCTGTGTTGTTATGTTTATGTTTAAACATTTATTAG 1379
Qy 1381 GTACTATCTATGAATGATTTTAAATATTTTTCATATTTCTGTGACAAAGCATTTTAAATTTG 1440
Db 1380 GTACTATCTATGAATGATTTTAAATATTTTTCATATTTCTGTGACAAAGCATTTTAAATTTG 1439
Qy 1441 CAAACAGTGGAGTCCATTTAGCCAGTGGGAAGCTTGGAACTCAGGTACCCCTTGAG 1500
Db 1440 CAAACAGTGGAGTCCATTTAGCCAGTGGGAAGCTTGGAACTCAGGTACCCCTTGAG 1499
Qy 1501 GATATGCTGCAGCCATCTCTTTGATCTGTCTTAACTGTAATTTATAGACAGCTAAA 1560
Db 1500 GATATGCTGCAGCCATCTCTTTGATCTGTCTTAACTGTAATTTATAGACAGCTAAA 1559
Qy 1561 TCCTTAACCTTGGATCTGGAAATGATTTAGACCTTGATACATTTCCAGAAATTTTCAGG 1620
Db 1560 TCCTTAACCTTGGATCTGGAAATGATTTAGACCTTGATACATTTCCAGAAATTTTCAGG 1619
Qy 1621 GGCATCGTGGTGTGGTCTAGTCTGATTTGAAAAACACAGAAACAGAGATCCAGCTGAAAA 1680
Db 1620 GGCATCGTGGTGTGGTCTAGTCTGATTTGAAAAACACAGAAACAGAGATCCAGCTGAAAA 1679
Qy 1681 GAGTGATCTCAATATCTCAATCTAATCTGCTCCTCAACTCAAGCAGAGTTTCTTCACTGTG 1740
Db 1680 GAGTGATCTCAATATCTCAATCTAATCTGCTCCTCAACTCAAGCAGAGTTTCTTCACTGTG 1739
Qy 1741 GCATGTCATCATCAACATTTAGTAGAGGATTTGTGATTTTATACAAATTTTAAATACA 1800
Db 1740 GCATGTCATCATCAACATTTAGTAGAGGATTTGTGATTTTATACAAATTTTAAATACA 1799
Qy 1801 ATGTCTTACATTTGATAAAAATCTTAAAGAGCAAAACTGCAATTTTCTGCAATCCACAT 1860
Db 1800 ATGTCTTACATTTGATAAAAATCTTAAAGAGCAAAACTGCAATTTTCTGCAATCCACAT 1859
Qy 1861 TCCATCATATTAGAATAGATATTTATCTATGAGATATAATGTCGAGAGACTT 1920
Db 1860 TCCATCATATTAGAATAGATATTTATCTATGAGATATAATGTCGAGAGACTT 1919
Qy 1921 TCATCTGTGATTCGCTGTTGTTTCTTAGGGTTCTTAGCCTGTAGCCTGACAGCATGTG 1980
Db 1920 TCATCTGTGATTCGCTGTTGTTTCTTAGGGTTCTTAGCCTGTAGCCTGACAGCATGTG 1979

Qy 1981 ATATGTGAATAAATGGATTCCTCTATAGCTAAATGATTCCTCTCGGGAGAGTTCTG 2040
Db 1980 ATATGTGAATAAATGGATTCCTCTATAGCTAAATGATTCCTCTCGGGAGAGTTCTG 2039
Qy 2041 GTACTCAATCACAATGCCAGATGGTTTATGCGCTATTTGTTGTAAGTAAGTGGTAAGA 2100
Db 2040 GTACTCAATCACAATGCCAGATGGTTTATGCGCTATTTGTTGTAAGTAAGTGGTAAGA 2099
Qy 2101 TGCATGAAGTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2160
Db 2100 TGCATGAAGTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2159
Qy 2161 TATGGAATAAATTTTGGTGCATATGATGTCATTCATTCATTCATTCATTCATTCATTCAT 2220
Db 2160 TATGGAATAAATTTTGGTGCATATGATGTCATTCATTCATTCATTCATTCATTCATTCAT 2219
Qy 2221 TTGTATTTATGAT 2280
Db 2220 TTGTATTTATGAT 2279
Qy 2281 TTTTGTACATATTTTACTTTGAAAAATATTTTAAATGGAAATTTTAAATPAAACATTTGATGT 2340
Db 2280 TTTTGTACATATTTTACTTTGAAAAATATTTTAAATGGAAATTTTAAATPAAACATTTGATGT 2339
Qy 2341 TTACAT 2346
Db 2340 TTACAT 2345

RESULT 3

AAS94792

ID AAS94792 standard; DNA; 2321 BP.

XX AAS94792;

XX 14-FEB-2002 (first entry)

XX Human DNA sequence #47 expressed during foam cell differentiation.

XX Human; foam cell differentiation; atherosclerosis; cerebral stroke;
XX cardiovascular disorder; coronary artery disease; gene therapy; ds.

XX Homo sapiens.

XX WO200177389-A2.

XX 18-OCT-2001.

XX 04-APR-2001; 2001WO-US011128.

XX 05-APR-2000; 2000US-0195106P.

XX (INCY-) INCYTE GENOMICS INC.

XX Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;
PI Tai J;

XX WPI; 2002-010925/01.

XX Composition useful for diagnosis of conditions, disorders or diseases
XX associated with atherosclerosis, comprises several polynucleotides that
XX are differentially expressed in foam cell development.

XX Claim 1; Page 102-103; 315pp; English.

XX The present invention relates to the isolation of human polynucleotide
XX sequences that are differentially expressed during foam cell
XX differentiation. The polynucleotide sequences of the invention or a
XX composition comprising these polynucleotides are useful as a high
XX throughput method for detecting altered expression of one or more
XX polynucleotides in a sample. The polynucleotides can be used in the
XX diagnosis of disorders associated with foam cell development such as

CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as
CC coronary artery disease. The polynucleotide sequences can also be used as
CC PCR primers and probes. The polynucleotides of the invention are also
CC useful in gene therapy. AAS94746-AAS95021 represent the human
CC polynucleotide sequences of the invention which are differentially
CC expressed during foam cell differentiation
XX
SQ Sequence 2321 BP; 700 A; 358 C; 421 G; 841 T; 0 U; 1 Other;

Query Match 98.2%; Score 2304.4; DB 6; Length 2321;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2316; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 28 CAGGTATCTCTTTTGCAGTGGTCTGATTTTGGTACCTTGTGAGTACCTTGTGACAGGGTATTAT 87
DB 4 CAGGTATCTCTTTTGCAGTGGTCTGATTTTGGTACCTTGTGAGTACCTTGTGACAGGGTATTAT 62
QY 88 TACAGCATCTTGGGAAACCTATTAGCCCTTCGATGTTTAAAGTATTAAGCTGTATAATTTGTTG 147
DB 63 TACAGCATCTTGGGAAACCTATTAGCCCTTTCGATGTTTAAAGTATTAAGCTGTATAATTTGTTG 122
QY 148 GGTGTGAGTGGTCTGACCTTAAATGTATTAATAAATTTAGACATCAAAATTTTCTCTACT 207
DB 123 GGTGTGAGTGGTCTGACCTTAAATGTATTAATAAATTTAGACATCAAAATTTTCTCTACT 182
QY 208 AACTAATCTTATAGTATGATCTTGGAGCAGTATATATATATATATATATATATATATATATAT 267
DB 183 AACTAATCTTATAGTATGATCTTGGAGCAGTATATATATATATATATATATATATATATATAT 242
QY 268 ATGTGTTACCTGGTCTAGGTTTGAACCTGCTTATTTTCAAAAGATTTTCTGAATTAATTT 327
DB 243 ATGTGTTACCTGGTCTAGGTTTGAACCTGCTTATTTTCAAAAGATTTTCTGAATTAATTT 302
QY 328 TTCCCTAGAAATTTCTCTTCAATCCAAAGTCAAAATATATATATATATATATATATATATATAT 387
DB 303 TTCCCTAGAAATTTCTCTTCAATCCAAAGTCAAAATATATATATATATATATATATATATAT 362
QY 388 GTTCCCAATGATGTATGCTCATPACTCGACTGAAACGATCTATGTTAAATGACTGTGTAT 447
DB 363 GTTCCCAATGATGTATGCTCATPACTCGACTGAAACGATCTATGTTAAATGACTGTGTAT 422
QY 448 ATGAATATTTCAAGTACTACCCCAAT 507
DB 423 ATGAATATTTCAAGTACTACCCCAAT 482
QY 508 ATGTAATCACTATGATTTATGCAAAACCAAGTATCTCCAAATTTTAAAGTAAATC 567
DB 483 ATGTAATCACTATGATTTATGCAAAACCAAGTATCTCCAAATTTTAAAGTAAATC 542
QY 568 TGATCTCTCTTGGAGAAATTTGATCTATATATATATATATATATATATATATATATATATATAT 627
DB 543 TGATCTCTCTTGGAGAAATTTGATCTATATATATATATATATATATATATATATATATATATAT 602
QY 628 CTGGTATAAATCAAAATTTCTATAAATTTCTACTTAACTCTTAAGAACTGGGTTCTTCTCT 687
DB 603 CTGGTATAAATCAAAATTTCTATAAATTTCTACTTAACTCTTAAGAACTGGGTTCTTCTCT 662
QY 688 TTGATGTTATTCATGTTTCAAGAGGAAACCAACATTTTACTCTTTTGGAGCAATTCCTAGA 747
DB 663 TTGATGTTATTCATGTTTCAAGAGGAAACCAACATTTTACTCTTTTGGAGCAATTCCTAGA 722
QY 748 ATCTATAGTATGATCAGGATATATTTGCTTTTAAATATATATTTTGGTTATTTTGAATACA 807
DB 723 ATCTATAGTATGATCAGGATATATTTGCTTTTAAATATATATTTTGGTTATTTTGAATACA 782
QY 808 GACATTTGGCTCCAAATTTTCTCTTTCGACATAGTATGCTTTTCTACTAGAACTTCTCA 867
DB 783 GACATTTGGCTCCAAATTTTCTCTTTCGACATAGTATGCTTTTCTACTAGAACTTCTCA 842
QY 868 ACATTTGGAACTTTGCAAAATATGAGCATCATATGTTTAAAGGCTGTATCTATTTAATGCT 927
DB 843 ACATTTGGAACTTTGCAAAATATGAGCATCATATGTTTAAAGGCTGTATCTATTTAATGCT 902

QY 928 ATGAGATACATTTGTTTCTCCTATGCTCAACACAGGTGAACAAACGCTAGTTGTTTTTACT 987
DB 903 ATGAGATACATTTGTTTCTCCTATGCTCAACACAGGTGAACAAACGCTAGTTGTTTTTACT 962
QY 988 GATATCTAAATGTTGGCTACCTGTGATTTTATAGTATGACATGTGACAAAAGCAAGAC 1047
DB 963 GATATCTAAATGTTGGCTACCTGTGATTTTATAGTATGACATGTGACAAAAGCAAGAC 1022
QY 1048 AATGGGCTCTTGACTGATCTGCAATCTTGGCAAACTTATGGTCTTCAATTTCTGACAGAC 1107
DB 1023 AATGGGCTCTTGACTGATCTGCAATCTTGGCAAACTTATGGTCTTCAATTTCTGACAGAC 1082
QY 1108 AGGATTTGACTCAATATTTTGTAGAGCTTGGCTAGAGTGAATACATGCTAGTGTGACT 1167
DB 1083 AGGATTTGACTCAATATTTTGTAGAGCTTGGCTAGAGTGAATACATGCTAGTGTGACT 1142
QY 1168 GGTAGAAATGTTTTAGTATTTGACTTCAGAAATTCATCTCAGGATGAATCTTTTATGCT 1227
DB 1143 GGTAGAAATGTTTTAGTATTTGACTTCAGAAATTCATCTCAGGATGAATCTTTTATGCT 1202
QY 1228 TTTTATTTGAAGCATATCTGAATTTTACTTTTATAAGATGGTTTTAGAAAGCTTTTGTCTAA 1287
DB 1203 TTTTATTTGAAGCATATCTGAATTTTACTTTTATAAGATGGTTTTAGAAAGCTTTTGTCTAA 1262
QY 1288 AAATTTGGCTTAGGATGTTAACTTCAATTTTCAGTTCGCAAGGGGTAGAAAATAATG 1347
DB 1263 AAATTTGGCTTAGGATGTTAACTTCAATTTTCAGTTCGCAAGGGGTAGAAAATAATG 1322
QY 1348 TGTGTTGTTATTTTATGTTAAACATATTTATTTAGGTACTATCTATGAATGTATTTAAATAT 1407
DB 1323 TGTGTTGTTATTTTATGTTAAACATATTTATTTAGGTACTATCTATGAATGTATTTAAATAT 1382
QY 1408 TTTTTCATATTTCTGTGACAGCATTTTAAATTTGCAACAGTGGAGTCCATTTAGCCAGT 1467
DB 1383 TTTTTCATATTTCTGTGACAGCATTTTAAATTTGCAACAGTGGAGTCCATTTAGCCAGT 1442
QY 1468 GGGAAAGTCTTGGAACTCAGGTTACCTTTGAAGGATATGCTGGCAGCCATCTCTTTGATC 1527
DB 1443 GGGAAAGTCTTGGAACTCAGGTTACCTTTGAAGGATATGCTGGCAGCCATCTCTTTGATC 1502
QY 1528 TGTGTTTAACTGTTAAATTTTATAGACAGTAAATCCCTTAATCTTGGATCGGAATGCAATTA 1587
DB 1503 TGTGTTTAACTGTTAAATTTTATAGACAGTAAATCCCTTAATCTTGGATCGGAATGCAATTA 1562
QY 1588 GTTATGACCTTTGTACCAATTTCCAGAAATTTCAAGGAGTATGCTGGGCTTTGGTCTAGTATTG 1647
DB 1563 GTTATGACCTTTGTACCAATTTCCAGAAATTTCAAGGAGTATGCTGGGCTTTGGTCTAGTATTG 1622
QY 1648 AAACACAGAGACAGAGATCCAGCTGAAAGAGAGTATCCCTCAATATCTTAACT 1707
DB 1623 AAACACAGAGACAGAGATCCAGCTGAAAGAGAGTATCCCTCAATATCTTAACT 1682
QY 1708 GGTCTCAACTCAAGCAGAGTTCTTCACTCTGGCACTGTGATCATGAAACTTAGTAGAG 1767
DB 1683 GGTCTCAACTCAAGCAGAGTTCTTCACTCTGGCACTGTGATCATGAAACTTAGTAGAG 1742
QY 1768 GGGATTTGCTGTTATTTATACAAATTTTAAATACATGTTTACATTTGATAAATTTCTTAA 1827
DB 1743 GGGATTTGCTGTTATTTATACAAATTTTAAATACATGTTTACATTTGATAAATTTCTTAA 1802
QY 1828 GAGCAAACTGCAATTTTATTTCTGATCCCAATTTCCCAATCATATTTAGAACTAAGATATTT 1887
DB 1803 GAGCAAACTGCAATTTTATTTCTGATCCCAATTTCCCAATCATATTTAGAACTAAGATATTT 1862
QY 1888 ATCTATGAGATATAAATTTGTTGACAGAGACTTTTCACTCTGGATGCTGTTTCTTAG 1947
DB 1863 ATCTATGAGATATAAATTTGTTGACAGAGACTTTTCACTCTGGATGCTGTTTCTTAG 1922
QY 1948 GGTCTCTAGCTGATGCTGACCAAGCATGTGATATGTGAAATATAAATGGAATCTTCTTA 2007
DB 1923 GGTCTCTAGCTGATGCTGACCAAGCATGTGATATGTGAAATATAAATGGAATCTTCTTA 1982
QY 2008 TAGCTAAATAGTTCTTCTCTGGGAGAGATTTCTGCTACTGCAATCAACAATCCAGATGGTG 2067

Db 1983 TAGCTAAATGAGTCCCTCTGGGAGAGTCTGGTACTGCAATCAATGCCAGATGGTG 2042
 Qy 2068 TTTATGGCTATTGTTGTAAGTAAAGTGTAAAGTCTATGAAGTAAAGTGTGTTGTTTC 2127
 Db 2043 TTTATGGCTATTGTTGTAAGTAAAGTGTAAAGTCTATGAAGTAAAGTGTGTTGTTTC 2102
 Qy 2128 ATCTTATGGAACCTCTTGATGATGCTGCTTTGATGGAATAAATTTTGGTGAATATCA 2187
 Db 2103 ATCTTATGGAACCTCTTGATGATGCTGCTTTGATGGAATAAATTTTGGTGAATATCA 2162
 Qy 2188 TGTCATTCACCTTTGCAATGAAATTTGGTTGTTGTTATTTATGTTATTTACCTGTCA 2247
 Db 2163 TGTCATTCACCTTTGCAATGAAATTTGGTTGTTGTTATTTATGTTATTTACCTGTCA 2222
 Qy 2248 GCTCTTCTAGTGTCTCAACCTTTTATAACCAATTTTGTACATATTTTACTTGAAATAT 2307
 Db 2223 GCTCTTCTAGTGTCTCAACCTTTTATAACCAATTTTGTACATATTTTACTTGAAATAT 2282
 Qy 2308 TTTAAATGGAATTTAAATAAACATTTTGATAGTTTACAT 2346
 Db 2283 TTTAAATGGAATTTAAATAAACATTTTGATAGTTTACAT 2321

RESULT 4

ID AAA98844 standard; cDNA; 1626 BP.
 AC AAA98844;

16-FEB-2001 (first entry)

Human differentiation-associated endothelial EST cDNA SEQ ID NO 20.

EST; expressed sequence tag; human; cell differentiation; antidiabetic;
 cell proliferation; endothelial tissue; angiogenic; antipsoriatic;
 ophthalmological; nephrotropic; cytostatic; hepatotropic; antiarthritic;
 antiarteriosclerotic; antirheumatic; gene therapy; angiogenic disease;
 treatment; rheumatoid arthritis; haemangioma; angiofibroma; eye disease;
 diabetic retinopathy; glaucoma; kidney disease; glomerulonephritis;
 diabetic nephropathy; malign nephrosclerosis; transplant rejection;
 thrombotic microangiopathic syndrome; glomerulopathy; fibrotic disease;
 mesangial cell proliferative disease; ss.

Homo sapiens.

DE19911684-A1.

14-SEP-2000.

09-MAR-1999; 99DE-01011684.

09-MAR-1999; 99DE-01011684.

01-OCT-1999; 99DE-01048679.

(SCHD) SCHERING AG.

Glienke J, Thierauch K, Hinzmann B, Pilarsky C;

WPI; 2000-572267/53.

Nucleic acid sequences from human endothelial cells, useful for gene
 therapy of angiogenesis and for identifying antiangiogenic agents.

Claim 1a; Page 41-42; 84pp; German.

This invention describes novel human nucleic acid sequences (I) from
 endothelial cells which have angiogenic, antipsoriatic, antidiabetic,
 ophthalmological, nephrotropic, cytostatic, hepatotropic,
 antiarteriosclerotic, antiarthritic, antirheumatic activity and can be
 used for gene therapy. The nucleic acid sequences may find use as tools
 to identify agents against angiogenic diseases. The polypeptides may be
 used in medicaments for gene therapy to treat angiogenic diseases. (I) in

CC sense or antisense form may be used. The genomic genes may also be used
 CC with suitable promoters and/or enhancers. The sequences (nucleic acid and
 CC proteins) may be used to treat psoriasis, arthritis, especially
 CC rheumatoid arthritis, haemangioma, angiofibroma, eye diseases, especially
 CC diabetic retinopathy, neovascular glaucoma, kidney diseases, such as
 CC glomerulonephritis, diabetic nephropathy, malign nephrosclerosis,
 CC thrombotic microangiopathic syndrome, transplantation rejection and
 CC glomerulopathy, fibrotic disease, such as liver cirrhosis, mesangial cell
 CC proliferative disease, arteriosclerosis and injury to nerve tissue
 XX

SQ Sequence 1626 BP; 493 A; 239 C; 306 G; 588 T; 0 U; 0 Other;

Query Match 58.2%; Score 1599.4; DB 3; Length 1626;

Best Local Similarity 59.3%; Pred. No. C;

Matches 161; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 734 GGCAATTCCTAGAACTCTATAGTAGTATCAGGATATATTTGCTTTAAATATATTTTGG 793

Db 1 GGCAATTTCTAGAACTCTATAGTAGTATCAGGATATATTTGCTTTAAATATATTTTGG 60

Qy 794 TTATTTTGAATACAGACATTTGGCTCCAAATTTTCTCTTCGACATAGTATGCTTTTC 853

Db 61 TTATTTTGAATACAGACATTTGGCTCCAAATTTTCTCTTCGACATAGTATGCTTTTC 120

Qy 854 ACTAGAACTTTCTCAACATTTGGAACTTTGCAAAATATGAGCATCATATGTTTAAAGGCTG 913

Db 121 ACTAGAACTTTCTCAACATTTGGAACTTTGCAAAATATGAGCATCATATGTTTAAAGGCTG 180

Qy 914 TATCATTTAATGCTATGAGATACATTTGTTCTCCCTATGCCAACAGGTGAACAAAGT 973

Db 181 TATCATTTAATGCTATGAGATACATTTGTTCTCCCTATGCCAACAGGTGAACAAAGT 240

Qy 974 AGTTGTTTTTACTGATACATAAATGTTGGCTACCTGTGATTTATGATATGACATGTCAT 1033

Db 241 AGTTGTTTTTACTGATACATAAATGTTGGCTACCTGTGATTTATGATATGACATGTCAT 300

Qy 1034 GAAAGGCAAGCAAACTGCTCTTGCTACTCAATCTCGGCAAACTTATTTGGTCTTTC 1093

Db 301 GAAAGGCAAGCAAACTGCTCTTGCTACTCAATCTCGGCAAACTTATTTGGTCTTTC 360

Qy 1094 ATTTCTCTGACAGACAGATTTGATCAATATTTGTTAGAGTTCCTGAATGGAATATCAT 1153

Db 361 ATTTCTCTGACAGACAGATTTGATCAATATTTGTTAGAGTTCCTGAATGGAATATCAT 420

Qy 1154 GGTAGTGATGCACTGGTGAAGATGTTTTAGTTATTCAGCTCAGAAATTCATCTCAGGATG 1213

Db 421 GGTAGTGATGCACTGGTGAAGATGTTTTAGTTATTCAGCTCAGAAATTCATCTCAGGATG 480

Qy 1214 AATCTTTTATGCTTTTATTTGTTAGAGCATATCTGAATTTACTTTTAAAGATGTTTATG 1273

Db 481 AATCTTTTATGCTTTTATTTTATGTTAGAGCATATCTGAATTTACTTTTAAAGATGTTTATG 540

Qy 1274 AAAGCTTTGCTTAAATAATTTGGCTTAGGAATGTTAACTTTTCACTTTTCCCAAGGGGT 1333

Db 541 AAAGCTTTGCTTAAATAATTTGGCTTAGGAATGTTAACTTTTCACTTTTCCCAAGGGGT 600

Qy 1334 AGAAATAATATGTTGTTGTTATTTGTTTATGTTTAACTATATATAGTACTATCTATGA 1393

Db 601 AGAAATAATATGTTGTTGTTATTTTATGTTTAACTATATATAGTACTATCTATGA 660

Qy 1394 ATGTATTTAAATATTTTTCATATTTCTGTGCAAGCATTTATAATTTGCAACAAGTGAAGT 1453

Db 661 ATGTATTTAAATATTTTTCATATTTCTGTGCAAGCATTTATAATTTGCAACAAGTGAAGT 720

Qy 1454 CCATTTAGCCAGTGGGAAGTCTTGGAACTCAGTTACCTTGAAGATATGCTGGCAG 1513

Db 721 CCATTTAGCCAGTGGGAAGTCTTGGAACTCAGTTACCTTGAAGATATGCTGGCAG 780

Qy 1514 CCATCTCTTTGATCTGTGCTTAACTGTAATTTATAGACCAGCTAAATCCCTAACTTGA 1573

Db 781 CCATCTCTTTGATCTGTGCTTAACTGTAATTTATAGACCAGCTAAATCCCTAACTTGA 840

Qy 1574 TCTGGAATGCAATAGTATGACCTTGTACCATTCCAGAAATTCAGGGGCACTCGTGGTT 1633

	Matches 1611;	Conservative	0;	Mismatches	1;	Indels	1;	Gaps	1;
QY	734	GGACAATTCCTAGAACTCTATAGTAGTATACAGGATATATTTTCGCTTTAAATAATATTTTGG	793						
Db	1	GGACAATTCCTAGAACTCTATAGTAGTATACAGGATATATTTTCGCTTTAAATAATATTTTGG	60						
QY	794	TTATTTTGAATACAGACATTTGGCTCCAAATTTTTCATCTTTGCACAAATAGTAGATGACTTTTC	853						
Db	61	TTATTTTGAATACAGACATTTGGCTCCAAATTTTTCATCTTTGCACAAATAGTAGATGACTTTTC	120						
QY	854	ACTAGAACTTCTCAACATTTTGGGAACCTTTGGCAATATGAGCATCATATGTGTAAAGCGTG	913						
Db	121	ACTAGAACTTCTCAACATTTTGGGAACCTTTGGCAATATGAGCATCATATGTGTAAAGCGTG	180						
QY	914	TATCATTTAAATGCTATGAGATACATCTTTTCTCCCTATGCCAAACAGGTGAACAAACGT	973						
Db	181	TATCATTTAAATGCTATGAGATACATCTTTTCTCCCTATGCCAAACAGGTGAACAAACGT	240						
QY	974	AGTTGTTTTTACTGATACATAAATGTGGCTACTCGTGTATTTTATAGTAGTAGACATGTCA	1033						
Db	241	AGTTGTTTTTACTGATACATAAATGTGGCTACTCGTGTATTTTATAGTAGTAGACATGTCA	300						
QY	1034	GAATAAGCAAGACAAATGGCTCTCTGACTGTAATACTTTGCGCAAACTTATTTGGGTCTTC	1093						
Db	301	GAATAAGCAAGACAAATGGCTCTCTGACTGTAATACTTTGCGCAAACTTATTTGGGTCTTC	360						
QY	1094	ATTTCCTGACAGACAGGAATTGACTCATATTTGTGTAGAGCTTGCCTAGAAATGGAATTACAT	1153						
Db	361	ATTTCCTGACAGACAGGAATTGACTCATATTTGTGTAGAGCTTGCCTAGAAATGGAATTACAT	420						
QY	1154	GGTAGTGTACACTGGTAGAAATGGTTTTTGTATTTTGTACTCAGAAATTCATCTCAGGATG	1213						
Db	421	GGTAGTGTACACTGGTAGAAATGGTTTTTGTATTTTGTACTCAGAAATTCATCTCAGGATG	480						
QY	1214	AATCTTTTATGTCCTTTTATGTAAAGCATATCTGAATTTTACTTTATTAAGATGGTTTTTAG	1273						
Db	481	AATCTTTTATGTCCTTTTATGTAAAGCATATCTGAATTTTACTTTTAAAGATGGTTTTTAG	540						
QY	1274	AAAGCTTTGTCTTAAAAATTTGGCCTAGGAATGGTAACTTTCATTTTCAGTTGCCAAGGGGT	1333						
Db	541	AAAGCTTTGTCTTAAAAATTTGGCCTAGGAATGGTAACTTTCATTTTCAGTTGCCAAGGGGT	600						
QY	1334	AGAAAAATTAATGTGTGTGTATGTTATGTTATCTATACATATTTATWAGTACTATCTATGA	1393						
Db	601	AGAAAAATTAATGTGTGTGTATGTTATGTTATCTATACATATTTATWAGTACTATCTATGA	660						
QY	1394	ATGTATTTAAATATTTTTCATATCTGTGACAAGCATTTTATAATTTGCAACAGTGGAGT	1453						
Db	661	ATGTATTTAAATATTTTTCATATCTGTGACAAGCATTTTATAATTTGCAACAGTGGAGT	720						
QY	1454	CCATTTAGCCAGTGGGAAGCTTTGGAATCTCAGGTTACCCCTGGAAGATATGCTGGCAG	1513						
Db	721	CCATTTAGCCAGTGGGAAGCTTTGGAATCTCAGGTTACCCCTGGAAGATATGCTGGCAG	780						
QY	1514	CCATCTCTTTGATCTGTGCTTAAACTGTAAATTTATAGACAGCTAAATCCCTTAACTTGG	1573						
Db	781	CCATCTCTTTGATCTGTGCTTAAACTGTAAATTTATAGACAGCTAAATCCCTTAACTTGG	840						
QY	1574	TCTGGAAATGCAATAGTTATGACTTTTACCATTTCCAGAAATTTCCAGGGGCATCTGGGTT	1633						
Db	841	TCTGGAAATGCAATAGTTATGACTTTTACCATTTCCAGAAATTTCCAGGGGCATCTGGGTT	899						
QY	1634	TGGTCTTAGTATTTGAAACACAAAGAACAGAGATCCAGCTGAAAAAGAGTGATCCTCAA	1693						
Db	900	TGGTCTTAGTATTTGAAACACAAAGAACAGAGATCCAGCTGAAAAAGAGTGATCCTCAA	959						
QY	1694	TATCTCTAACTACTGCTCCTCAACTCAAGCAGAGTTCCTTCTCTGGCAGCTGTGATCAT	1753						
Db	960	TATCTCTAACTACTGCTCCTCAACTCAAGCAGAGTTCCTTCTCTGGCAGCTGTGATCAT	1019						
QY	1754	GAACCTTAGTAGAGGGGATTTGTGTATTTTATACAAATTTAATCAATGTCTTACATTG	1813						
Db	1020	GAACCTTAGTAGAGGGGATTTGTGTATTTTATACAAATTTAATCAATGTCTTACATTG	1079						

RESULT 7	
AAH34083/c	
ID . AAH34083 standard; cDNA; 665 BP.	
XX	
AAH34083;	
XX	
AC	
XX	
DT	
XX	
03-SEP-2001 (first entry)	
XX	
DE	
XX	
Human colon cancer antigen encoding cDNA SEQ ID NO:1165.	
XX	
Human; colon cancer; colon cancer antigen; diagnosis; detection;	
KW	
XX	
colorectal carcinoma; ss.	
OS	
XX	
Homo sapiens.	
OS	
XX	
WG200122920-A2.	
PN	
XX	
05-APR-2001.	
XX	
28-SEP-2000; 2000WO-US026524.	
XX	
29-SEP-1999; 99US-0157137P.	
PR	
XX	
03-NOV-1999; 99US-0163280P.	
PR	
XX	
XX	
(HUMA-) HUMAN GENOME SCI INC.	
PA	
XX	
XX	
Ruben SM, Barash SC, Birse CE, Rosen CA;	
PI	
XX	
WPI: 2001-235357/24.	
DR	
XX	
P-PSDB; AAG74678.	
DR	
XX	
Nucleic acids encoding 4277 human colon cancer-associated polypeptide	
PT	
XX	
useful for preventing, diagnosing and/or treating colorectal cancer	
PT	
XX	
Claim 1; Page 2972; 9803pp; English.	
PS	

XX		AAH32943	to ANH37195	and AAG73514	to AAG77788	represent human colon
CC		cancer-associated nucleic acid molecules (N)	and proteins (P), where the			
CC		proteins are collectively known as colon cancer antigens. The colon				
CC		cancer antigens have cytostatic activity and can be used in gene therapy				
CC		and vaccine production. N and P may be used in the prevention, diagnosis				
CC		and treatment of diseases associated with inappropriate p expression. For				
CC		example, N and P may be used to treat disorders associated with decreased				
CC		expression by rectifying mutations or deletions in a patient's genome				
CC		that affect the activity of P by expressing inactive proteins or to				
CC		supplement the patients own production of P. Additionally, N may be used				
CC		to produce the colon cancer-associated ps, by inserting the nucleic acids				
CC		into a host cell and culturing the cell to express the proteins. N and P				
CC		can be used in the prevention, diagnosis and treatment of colorectal				
CC		carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent				
CC		sequences used in the exemplification of the present invention. N.B.				
CC		Pages 666 to 682 and page 7053 of the sequence listing were missing at				
CC		time of publication, meaning no sequences are present for SEQ ID NO:1027				
CC		cc to 1052, 7921 and 7922				
XX						
SQ		Sequence 665 BP; 234 A; 115 C; 84 G; 227 T; 0 U; 5 Other;				
	Query March	25.7%; Score 603.2; DB 4; Length 665;				
	Best Local Similarity	99.4%; Pred. No. 4.4e-118;				
	Matches 613; Conservative 2; Mismatches 1; Indels 1; Gaps 1;					
QY	1731	CTTCACACTCGCACCTGTCATCATGAACCTAGTAGAGGGGATTGTGTATTTTATACAA	1790			
DB	664	CTTCACACTCGCACCTGTCATCATGAACCTAGTAGAGGGGATTGTGTATTTTATACAA	605			
QY	1791	ATTTAAATACAAATGCTTACATGATAAATTTCTTTAAAGAGCAAACTGCATTTATTTCT	1850			
DB	604	ATTTAAATACAAATGCTTACATGATAAATTTCTTTAAAGAGCAAACTGCATTTATTTCT	545			
QY	1851	GCATCCCAATTCCTCAATCATATTAGAACAATAATTTATCTATGAAGATATAATGTGTC	1910			
DB	544	GCATCCCAATTCCTCAATCATATTAGAACAATAATTTATCTATGAAGATATAATGTGTC	485			
QY	1911	AGAGAGACTTTTCATCTGTCGGATTCGGTTGTTTCTTAGGGTTCCTAGCACTGATGCCCTGCA	1970			
DB	484	AGAGAGACTTTTCATCTGTCGGATTCGGTTGTTTCTTAGGGTTCCTAGCACTGATGCCCTGCA	425			
QY	1971	CAAGCATCTGATATGTGAAAATAAAATGGATTTCTTCTATAGCTAAATGAGTT-CCCTCTGG	2029			
DB	424	CAAGCATCTGATATGTGAAAATAAAATGGANTCTTCTATAGCTAAATGAGTTTCCTCTGG	365			
QY	2030	GGAGAGTTCTGGTACTCCAATCACAATCCAGATCGTGTTTATCGGCTATTGTGTAAGT	2089			
DB	364	GGAGAGTTCTGGTACTCCAATCACAATCCAGATCGTGTTTATCGGCTATTGTGTAAGT	305			
QY	2090	AAGTGGTAAGATGCTTGAAGTAAAGTGTGTTTGTGTTTTTCATCTTATGGAACCTCTGTATGC	2149			
DB	304	AAGTGGTAAGATGCTTGAAGTAAAGTGTGTTTGTGTTTTTCATCTTATGGAACCTCTGTATGC	245			
QY	2150	ATGTGCTTTTCTGATCGGAATAAATTTTGTGCGAATATGATGTCATTCACATTTGCATTGAA	2209			
DB	244	ATGTGCTTTTCTGATCGGAATAAATTTTGTGCGAATATGATGTCATTCACATTTGCATTGAA	185			
QY	2210	TTCGAATTTTGGTTGTATTTATATGATTAATACCCTCTACGCTTCCTAGTTGCTTCAACCAT	2269			
DB	184	TTCGAATTTTGGTTGTATTTATATGATTAATACCCTCTACGCTTCCTAGTTGCTTCAACCAT	125			
QY	2270	TTTATACCAATTTTGTACATATTTTACTTTGGAATAATTTTAAATGGAAAAATTTAAATAAA	2329			
DB	124	TTTATACCAATTTTGTACATATTTTACTTTGGAATAATTTTAAATGGAAAAATTTAAATAAA	65			
QY	2330	CATTGTGATAGTTTACAT	2346			
DB	64	CATTGTGATAGTTTACAT	48			

RESULT 8
ABZ08185

QY 708 AAAGGAAACACACCTTTACTCTTTTAGGACAACTTCTAGAACTCTATAGTATCAGAT 767
Db 241 AAAGGAAACACACCTTTACTCTTTTAGGACAACTTCTAGAACTCTATAGTATCAGAT 300
QY 768 ATATTTTGCTTTAAATATATTTTGGTTATTTGAAATACAGACATTTGGCTCCAAATTTTC 827
Db 301 ATATTTTGCTTTAAATATATTTTGGTTATTTGAAATACAGACATTTGGCTCCAAATTTTC 360
QY 828 ATCTTTGACAAATAGTATGACCTTTTCACTAGAACTTTCTCAACATTTTGGAACTTTGCAA 887
Db 361 ATCTTTGACAAATAGTATGACCTTTTCACTAGAACTTTCTCAACATTTTGGAACTTTGCAA 420
QY 888 TAGAGCAATCATATGTTTAAAGCTGTATCAATTTAAATGCTATGAGATATGTTTCTC 947
Db 421 TAGAGCAATCATATGTTTAAAGCTGTATCAATTTAAATGCTATGAGATATGTTTCTC 480
QY 948 CCTATGCCAAACAGGTGAACAAACCTAGTCTGTTTCTTACTGATATAAATGTTGGCTACC 1007
Db 481 CCTATGCCAAACAGGTGAACAAACCTAGTCTGTTTCTTACTGATATAAATGTTGGCTACC 540
QY 1008 TGTGATTTTATAGTATGACATGTGAGAAAGGCAAGCAAAATGGCTCTTG 1060
Db 541 TGTGATTTTATAGTATGACATGTGAGAAAGGCAAGCAAAATGGCTCTTG 593

RESULT 9

AAL22443
ID AAL22443 standard; cDNA; 608 BP.

XX AC AAL22443;

XX XX 07-DEC-2001 (first entry)

XX DE Human breast cancer expressed polynucleotide 14900.
XX KW Human; breast cancer; cell marker; cytostatic; ss.
XX OS Homo sapiens.

XX XX WO200151628-A2.

XX PD 19-JUL-2001.

XX PF 10-JAN-2001; 2001WO-US000799.

XX PR 14-JAN-2000; 2000US-0176077P.

XX PR 14-MAR-2000; 2000US-0189167P.

XX PR 24-MAR-2000; 2000US-0192099P.

XX PR 29-MAR-2000; 2000US-0193480P.

XX PR 15-MAY-2000; 2000US-0205230P.

XX PR 09-JUN-2000; 2000US-0211315P.

XX PR 25-JUL-2000; 2000US-0220534P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Lillie J. Xu Y, Wang Y, Steinmann K;

XX DR WPI; 2001-451856/48.

XX XX New peptide useful as a marker for the diagnosis of breast cancer.

XX XX Claim 1; Page 2691; 3695pp; English.

XX CC The invention relates to human breast cancer expressed polynucleotides
XX CC (AAL07544-AAL26789) and methods of assessing whether a patient is
XX CC afflicted with breast cancer by examining the correlation between the
XX CC expression of certain markers and the cancerous state of breast cells.
XX CC The polynucleotides and encoded polypeptides are potential markers for
XX CC detecting, diagnosing, and monitoring, characterising treating and
XX CC potentially preventing breast cancer. The polynucleotides and encoded
XX CC polypeptides are also useful for isolating compounds with cytostatic
XX CC activity

XX SQ Sequence 608 BP; 198 A; 101 C; 87 G; 222 T; 0 U; 0 Other;
Query Match 24.9%; Score 584.8; DB 4; Length 608;
Best Local Similarity 98.8%; Pred. No. 3.4e-114;
Matches 589; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 468 CCCAAATAAATCTTTCTTATTTGCTCTGAAAGAAAGAAAGCAATGTAATCACTATGATTAT 527
Db 10 CCCAAATAAATCTTTCTTATTTGCTCTGAAAGAAAGAAAGCAATGTAATCACTATGATTAT 69
QY 528 TGCACAAACACCAAGATTTCTCCAACTTTTAAAGTAATCTGATCCTCTTCTTGGAGAAA 587
Db 70 TGCACAAACACCAAGATTTCTCCAACTTTTAAAGTAATCTGATCCTCTTCTTGGAGAAA 129
QY 588 ATTTGTTACTAATAGTTTTTCTTATGAAATGTTTATTACTACTGCTATGTAATCAAAATTTCT 647
Db 130 ATTTGTTACTAATAGTTTTTCTTATGAAATGTTTATTACTACTGCTATGTAATCAAAATTTCT 189
QY 648 ATAAATTTCTTACTTAAAGTCTTAAAGTCTTAAAGTCTTAAAGTCTTAAAGTCTTAAAGTCT 707
Db 130 ATAAATTTCTTACTTAAAGTCTTAAAGTCTTAAAGTCTTAAAGTCTTAAAGTCTTAAAGTCT 249
QY 708 AAAGGAAACAACTTTTACTCTTTTAGGACAAATTTCTAGAACTCTATAGTATCAGAT 767
Db 250 AAAGGAAACAACTTTTACTCTTTTAGGACAAATTTCTAGAACTCTATAGTATCAGAT 309
QY 768 ATATTTTGCTTTAAATATATTTTGGTTATTTGAAATACAGACATTTGGCTCCAAATTTTC 827
Db 310 ATATTTTGCTTTAAATATATTTTGGTTATTTGAAATACAGACATTTGGCTCCAAATTTTC 369
QY 828 ATCTTTGCACAATAGTATGACTTTTCTACTAGAACTTTCTCAACATTTTGGAACTTTGCAA 887
Db 370 ATCTTTGCACAATAGTATGACTTTTCTACTAGAACTTTCTCAACATTTTGGAACTTTGCAA 429
QY 888 TATGAGCAATCATATGTTTAAAGCTGTATCAATTTAAATGCTATGAGATATGTTTCTC 947
Db 430 TATGAGCAATCATATGTTTAAAGCTGTATCAATTTAAATGCTATGAGATATGTTTCTC 489
QY 948 CCTATGCCAAACAGGTGAACAAACCTAGTCTGTTTCTTACTGATATAAATGTTGGCTACC 1007
Db 490 CCTATGCCAAACAGGTGAACAAACCTAGTCTGTTTCTTACTGATATAAATGTTGGCTACC 549
QY 1008 TGTGATTTTATAGTATGACATGTGAGAAAGGCAAGCAAAATGGCTCTTG 1063
Db 550 TGTGATTTTATAGTATGACATGTGAGAAAGGCAAGCAAAATGGCTCTTG 605

RESULT 10

ABK27756/c

ID ABK27756 standard; cDNA; 586 BP.

XX AC ABK27756;

XX DT 09-APR-2002 (first entry)

XX DE Human colon cancer expressed sequence tag, Seq ID no 193.

XX KW Human; colon cancer; T cell expansion; tumour; EST; gene; ss;

XX XX expressed sequence tag.

XX OS Homo sapiens.

XX PN WO200196390-A2.

XX PD 20-DEC-2001.

XX PF 08-JUN-2001; 2001WO-US018577.

XX PR 09-JUN-2000; 2000US-0210821P.

XX PR 18-DEC-2000; 2000US-0256571P.

XX PR 10-MAY-2001; 2001US-0250240P.

XX XX

(CORI-) CORIXA CORP.
Jiang Y, Hepler WT, Clapper JD, Wang A, Secrist H;
WPI; 2002-139708/18.
Novel isolated polynucleotide encoding a polypeptide comprising a portion
of colon tumor protein, useful for detection, diagnosis and therapy of
human colon cancer.
Claim 1; Page 202; 220pp; English.
The invention relates to an isolated polynucleotide (I) encoding a
polypeptide (II) comprising at least a portion of a colon tumour protein.
(I), (II) and antibody (III) to (II) are useful for determining the
presence of a cancer in a patient. (I), (II) or antigen presenting cells
expressing (I) is useful for stimulating and/or expanding T cells
specific for a tumour protein, by contacting T cells with (I), (II) or
antigen-presenting cells that express (I), under conditions and for a
time sufficient to permit the stimulation and/or expansion of T cells.
(I), (II) or antigen presenting cells that express (II) are useful for
treating colon cancer in a patient by incubating CD4+ and/or CD8+ T cells
isolated from a patient with (I), (II) or antigen presenting cells that
express (II), such that T cells proliferate, and administering to the
patient an effective amount of the proliferated T cells, thus inhibiting
the development of a cancer in the patient. (I) or (II) is useful in
vaccines and pharmaceutical compositions for prevention and treatment of
colon malignancies and for the diagnosis and monitoring of such cancers.
(I), (II) or (III) is useful for detection, diagnosis and/or therapy of
human colon cancer. (I) is useful as a probe or primer for nucleic acid
hybridisation, and in the design and preparation of ribozyme molecules
for inhibiting expression of (II) in tumour cells. ABK27564-ABK27807
represent novel human colon cancer coding sequences and primers of the
invention
XX
SQ
Sequence 586 BP; 214 A; 88 C; 97 G; 187 T; 0 U; 0 Other;
Query Match 24.8%; Score 582.8; DB 6; Length 586;
Best Local Similarity 99.7%; Pred. No. 8.9e-114;
Matches 584; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 476 AACTTCTTATTCCTCTGAAGAAGAAAGCAATGTAATCACTATGATTATTCACAAA 535
DB 586 AACTTCTTATTCCTCTGAAGAAGAAAGCAATGTAATCACTATGATTATTCGCGAAA 527
QY 536 CAACCAAGATCTCCAAACAAATTTAAAGTAATCTGATCCTCTTCTGGAGAAATTTGTTAC 595
DB 526 CAACCAAGATCTCCAAACAAATTTAAAGTAATCTGATCCTCTTCTGGAGAAATTTGTTAC 467
QY 596 CTAATAGTTTTTCTTATGATGATGTTATTACTACTGTTATATAATCAAAATTTCTATAAATTT 655
DB 466 CTAATAGTTTTTCTTATGATGATGTTATTACTACTGTTATATAATCAAAATTTCTATAAATTT 407
QY 656 CTTACTTAAGTCTTAAGAACTGGGTTCTTCTTTGATGTTATTTCATGTTTCAGAAAGGAAA 715
DB 406 CTTACTTAAGTCTTAAGAACTGGGTTCTTCTTTGATGTTATTTCATGTTTCAGAAAGGAAA 347
QY 716 CAACACTTTACTCTTTTAGCAATCTCTAGATCTATAGTATCAGATATATTTTG 775
DB 346 CAACACTTTACTCTTTTAGCAATCTCTAGATCTATAGTATCAGATATATTTTG 287
QY 776 CTTTAAATATATTTTGGTTATTTTGAATACAGACATTTGGCTCCAAATTTTCATCTTTGC 835
DB 286 CTTTAAATATATTTTGGTTATTTTGAATACAGACATTTGGCTCCAAATTTTCATCTTTGC 227
QY 836 ACAATAGTATGACTTTTTCATAGAACTTCTCAACATTTGGAACTTTGCAATATAGCA 895
DB 226 ACAATAGTATGACTTTTTCATAGAACTTCTCAACATTTGGAACTTTGCAATATAGCA 167
QY 896 TCATATGTTTAAAGCTGTATCATTTAATGCTATGAGATACATGTTTCTTCCCTATGCC 955
DB 166 TCATATGTTTAAAGCTGTATCATTTAATGCTATGAGATACATGTTTCTTCCCTATGCC 107

QY 956 AAACAGGTGAACAAACGCTAGTCTTTTCTTACTGATCTAAATGTTGGCTACCTGATTT 1015
DB 106 AAACAGGTGAACAAACGCTAGTCTTTTCTTACTGATCTAAATGTTGGCTACCTGATTT 47
QY 1016 TATAGTATGACATGTCAGAAAAAGGCAAGACAAATGGCTCTTTGT 1061
DB 46 TATAGTATGACATGTCAGAAAAAGGCAAGACAAATGGCTCTTTGT 1
RESULT 11
ABK39231
ID ABK39231 standard; cDNA; 675 BP.
XX
AC ABK39231;
XX
DT 21-MAY-2002 (first entry)
XX
DE DNA encoding lung tumour protein associated contig 89.
XX
KW Lung tumour; cancer; T cell; immune response stimulator; cytostatic;
contig; ss.
XX
OS Homo sapiens.
XX
FN WO200204514-A2.
XX
PD 17-JAN-2002.
XX
PP 10-JUL-2001; 2001WO-US022058.
XX
PR 11-JUL-2000; 2000US-00614124.
PR 29-AUG-2000; 2000US-00651563.
PR 08-SEP-2000; 2000US-00658824.
PR 26-SEP-2000; 2000US-00671325.
PR 06-OCT-2000; 2000US-00677419.
PR 30-OCT-2000; 2000US-00702705.
PR 13-DEC-2000; 2000US-00736457.
PR 03-MAY-2001; 2001US-00849626.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;
Manerakis M, Carter D, Fanger GR, Vedwick TS, Bangur CS, Mcnabb A;
PI Wang A, Fanger N, Switzer A, Mcneill PD, Clapper JD;
XX
DR WPI; 2002-164634/21.
XX
PT Novel polynucleotide encoding a lung tumor polypeptide useful for
stimulating and/or expanding T cells specific for a tumor protein.
XX
PS Claim 1; SEQ ID NO 1269; 223pp; English.
XX
CC The invention describes an isolated polynucleotide and polypeptide useful
for stimulating and/or expanding T cells specific for a tumour protein
for determining the presence of a cancer in a patient. A composition
containing the polynucleotide and/or polypeptide is useful for treating a
lung cancer in a patient. The polypeptide is useful for removing tumour
cells from a biological sample. The polynucleotide is also useful as
a probe or primer to detect the level of mRNA encoding a tumour protein.
CC This sequence represents a contig of a DNA sequence encoding a lung
tumour associated protein, described in the method of the invention.
CC Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 675 BP; 193 A; 110 C; 133 G; 236 T; 0 U; 3 Other;
Query Match 24.8%; Score 581.4; DB 6; Length 675;
Best Local Similarity 97.3%; Pred. No. 1.8e-113;
Matches 622; Conservative 0; Mismatches 13; Indels 4; Gaps 3;
QY 1673 CTGAAAAAGCTGATCTCTCAATATCTCAATGCTCTCACTCAAGCAGATTTCT 1732

Db 1 CTGAAAGAGTGATCCCTCAATATCCTAATCACTGATGCTCACTCAAGCAGAGTTTCT 60
Qy 1733 TCACCTCGGCACTGATGATCAAGAACTTAGTAGAGGGGATTTGTGTATTTATPACAAAT 1792
Db 61 TCACCTCGGCACTGATGATCAAGAACTTAGTAGAGGGGATTTGTGTATTTATPACAAAT 120
Qy 1793 TTAATPACAAATGCTTACATTCATGATAAATTTCTTAAAGAGCAAACTGCAATTTATTTCTGC 1852
Db 121 TTAATPACAAATGCTTACATTCATGATAAATTTCTTAAAGAGCAAACTGCAATTTATTTCTGC 180
Qy 1853 ATCCACATTCACATCATATTAAGACTAAGATATTTATCTATGAAGATATAAATGGTGCAG 1912
Db 181 ATCCACATTCACATCATATTAAGACTAAGATATTTATCTATGAAGATATAAATGGTGCAG 240
Qy 1913 AGAGACTTTTCATCTGCTGATTCGCTGTTTCTTCTAGGCTTCTAGCACTGATGCTGCACA 1972
Db 241 AGAGACTTTTCATCTGCTGATTCGCTGTTTCTTCTAGGCTTCTAGCACTGATGCTGCACA 300
Qy 1973 AGCATGTGATATGTAAGATAAATGATTTCTTATAGCTAAATGAGTTCCCTCTGGGA 2032
Db 301 AGCATGTGATATGTAAGATAAATGATTTCTTATAGCTAAATGAGTTCCCTCTGGGA 360
Qy 2033 GAGTTCTGGTACTGATCAATCAATGCCAGATGCTTTTATGGCTATTTGTTGTAAGTAAG 2092
Db 361 GAGTTCTGGTACTGATCAATCAATGCCAGATGCTTTTATGGCTATTTGTTGTAAGTAAG 420
Qy 2093 TGGTAAGATGCTATGAAGTATGTTGTTTCTTATGCTTATGAAACTCTTGTATGCAATG 2152
Db 421 TGGTAAGATGCTATGAAGTATGTTGTTTCTTATGCTTATGAAACTCTTGTATGCAATG 480
Qy 2153 TGCCTTTGATGGAATAAATTTGGTGAATATGATGATGATGATGATGATGATGATGATG 2212
Db 481 TGCCTTTGATGGAATAAATTTGGTGAATATGATGATGATGATGATGATGATGATGATG 540
Qy 2213 -AATTTTGGTGTATTTATGATATATA-CCTGTCAGCTTCTAGTTGCTTCAACCAAT 2270
Db 541 AATTTTGGTGTATTTATGATATATA-CCTGTCAGCTTCTAGTTGCTTCAACCAAT 600
Qy 2271 TTAATPACAAATTTGTCATATTTTACTTGAATAATTT 2309
Db 601 T--ATACCAATTTGNACATATTTTACTTGNAAATATTT 637

RESULT 12
ACAL1560
ID ACAL1560 standard; cDNA; 675 BP.
AC ACAL1560;
XX ACAL1560;

DT 05-JUN-2003 (first entry)
DE Human lung cancer library SQL1, cDNA SEQ ID 1269.

XX Human; lung cancer; ss; lung tumour; cytostatic; vaccine;
KW T cell expansion; CD4; CD8.
XX Homo sapiens.

XX US2002197669-A1.
XX 26-DEC-2002.

XX 03-MAY-2001; 2001US-00849626.
XX 13-DEC-2000; 2000US-00736457.

XX (BANG/) BANGUR C S.
PA (FANG/) FANGER G R.
PA (WANG/) WANG A.
PA (WANG/) WANG T.
PA (SWIT/) SWITZER A P.
PA (MCNE/) MCNEILL P D.
PA (CLAP/) CLAPPER J D.

XX Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;
PI Clapper JD;
XX
XX MPI; 2003-352750/33.

XX Novel lung cancer polynucleotide encoding lung cancer protein, useful for
PT detecting the presence of lung cancer in a patient, and in pharmaceutical
PT compositions, e.g. vaccines, for treating lung cancer.

XX Example 1; Page; 72pp; English.

XX The invention relates to a polynucleotide encoding a lung tumour protein,
CC comprising a sequence selected from any of the 14 sequences mentioned in
CC the specification, or a sequence (S2) mentioned in the specification,
CC complement of S1, sequences consisting of at least 20 contiguous residues
CC of S1, sequences that hybridise to S1, sequences having 75% preferably
CC 90% identity to S1, or degenerate variants of S1. Also included are an
CC isolated polypeptide (comprising a sequence (S3) selected from any one of
CC the 4 amino acid sequences mentioned in the specification, a sequence
CC encoded by the polynucleotide, or sequences having at least 70%,
CC preferably 90% identity to a sequence encoded by the polynucleotide), an
CC expression vector comprising the polynucleotide operably linked to an
CC expression control sequence, a host cell transformed or transfected with
CC the vector, an isolated antibody (or its antigen-binding fragment) that
CC specifically binds to the polypeptide, detecting the presence of a cancer
CC oligonucleotide that hybridises to S1 under moderately stringent
CC conditions, stimulating and/or expanding T cells with the polynucleotide, a
CC protein (comprising contacting T cells with the polynucleotide, protein
CC or antigen-presenting cells, under conditions and for a time sufficient
CC to permit the stimulation and/or expansion of T cells) and inhibiting the
CC development of a cancer in a patient (by incubating CD4+ and/or CD8+ T
CC cells isolated from a patient with the polynucleotide, protein or antigen
CC presenting cells that express the polynucleotide, such that T cells
CC proliferate, administering to the patient an effective amount of the
CC proliferate T cells, and thus inhibiting the development of a cancer in
CC the patient. The polynucleotide, protein and cells are useful in a
CC composition for stimulating an immune response in a patient, and for
CC treating a cancer in a patient (particularly lung cancer). The
CC oligonucleotide is useful for determining the presence of a cancer in a
CC patient. The protein and oligonucleotides are useful in pharmaceutical
CC compositions, e.g. vaccines. The polynucleotide is also useful as a probe
CC or primer for nucleic acid hybridisation, and in the design and
CC preparation of ribozyme molecules for inhibiting expression of tumour
CC polypeptides and proteins in tumour cells. An amplified portion of the
CC polynucleotide is useful for isolating a full-length gene from a suitable
CC library. The present sequence is a cDNA (full length, extended or
CC partial) isolated from a library derived from lung tumour/cancer cells.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO at seqdata.uspto.gov/sequence.html?DocId=20020197669

XX Sequence 675 BP; 193 A; 110 C; 133 G; 236 T; 0 U; 3 Other;

Query Match 24.8%; Score 581.4; DB 7; Length 675;
Best Local Similarity 97.3%; Pred. No. 1.8e-113;
Matches 622; Conservative 0; Mismatches 13; Indels 4; Gaps 3;

Qy 1673 CTGAAAAGAGTGATCCTCAATATCCTAATCACTGATGCTCCTCACTCAAGCAGAGTTTCT 1732
Db 1 CTGAAAAGAGTGATCCTCAATATCCTAATCACTGATGCTCCTCACTCAAGCAGAGTTTCT 60
Qy 1733 TCACCTCGGCACTGATGATCAAGAACTTAGTAGAGGGGATTTGTGTATTTATPACAAAT 1792
Db 61 TCACCTCGGCACTGATGATCAAGAACTTAGTAGAGGGGATTTGTGTATTTATPACAAAT 120
Qy 1793 TTAATPACAAATGCTTACATTCATGATAAATTTCTTAAAGAGCAAACTGCAATTTATTTCTGC 1852
Db 121 TTAATPACAAATGCTTACATTCATGATAAATTTCTTAAAGAGCAAACTGCAATTTATTTCTGC 180
Qy 1853 ATCCACATTCACATCATATTAAGACTAAGATATTTATCTATGAAGATATAAATGGTGCAG 1912

Human colon tumour antigen polynucleotide SEQ ID NO:1423.
Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
colon tumour metastatic antigen; diagnosis; gene; ss.

Homo sapiens.

WO200196388-A2.

20-DEC-2001.

08-JUN-2001; 2001WO-US018557.

09-JUN-2000; 2000US-0210899P.

20-FEB-2001; 2001US-0270216P.

(CORI-) CORIXA CORP.

Jiang Y, Harlocker SL, Secret H;

WPI; 2002-114514/15.

Novel isolated colon tumor polynucleotide differentially expressed in colon tumor or colon metastatic tumor and polypeptides encoded by them, useful for inhibiting development of cancer in patient.

Claim 1; SEQ ID NO 1423; 105pp; English.

ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I) which were isolated from human colon tumour and colon metastatic tumour cDNA libraries. (I) have cytostatic activity and can be used in vaccine production. (I) can be used for stimulating and/or expanding T cells specific for a tumour protein on contact with the T cells. They are also useful for inhibiting the development of cancer in a patient. (I) can be used as probes or primers for nucleic acid hybridisation, for preparing mutant species primers, or primers for use in genetic constructions. (I) can be used in the diagnosis of a colon tumour

Sequence 586 BP; 189 A; 98 C; 85 G; 211 T; 0 U; 3 Other;

Query Match 24.2%; Score 588; DB 6; Length 586;

Best Local Similarity 99.0%; Pred. No. 1.2e-110;

Matches 580; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

464 ACTACCCCAATAAATCTTCTTATGCTCTGGAAGAGAAAGCAATGTAATCACTATGA 523

1 ACTACCCCAATAAATCTTCTTATGCTCTGGAAGAGAAAGCAATGTAATCACTATGA 60

524 TTATGCGCAACCAACAGAAATCTCCAAACAATTTTAAAGTAATCTGATCCTCTTCTTGA 583

61 TTATGCGCAACCAACAGAAATCTCCAAACAATTTTAAAGTAATCTGATCCTCTTCTTGA 120

584 GAAATGTTACTTAATAGTCTTTTCTTATGATGTTTACTACTGGTATAAATCAAT 643

121 GAAATGTTACTTAATAGTCTTTTCTTATGATGTTTACTACTGGTATAAATCAAT 180

644 TTCTATAAATTTCTTACTTAAGTCTTAAGAACTGGGTTCTTCTTGTATGTTATTCATGT 703

181 TTCTATAAATTTCTTACTTAAGTCTTAAGAACTGGGTTCTTCTTGTATGTTATTCATGT 240

704 TCAGAAGGAACAACACATTTACTCTTTTGAACAATTCCTAGAACTCTATAGTATCA 763

241 TCAGAAGGAACAACACATTTACTCTTTTGAACAATTCCTAGAACTCTATAGTATCA 300

764 GGATATATTTTGGTTTAAATATATTTTGGTTATTTTGAATACAGACATTTGGCTCCAAT 823

301 GGATATATTTTGGTTTAAATATATTTTGGTTATTTTGAATACAGACATTTGGCTCCAAT 360

824 TTTCATCTTTGCAATAGTATGACTTTTCACTAGAACTTCTCAACATTTGGGAATTTG 883

361 TTTCATCTTTGCAATAGTATGACTTTTCACTAGAACTTCTCAACATTTGGGAATTTG 420

884 CAAATATGAGCATCATATCTGTTAAGGTG-TATCATTTTAATGCTATGAGATACATTGT 942

421 CAAATATGAGCATCATATCTGTTAAGGTGTTATCATTTAATGCTATGAGATACATTGT 480

943 TTCTCCCTATGCCAACAACAGGTGAACAACAGTGTGTTTCTTACTGATATAAATCTTGG 1002

481 TTCTCCCTATGCCAACAACAGGTGAACAACAGTGTGTTTCTTACTGATATAAATCTTGG 540

1003 CTACCTGTGATTTTATGATGACATGTGAGAAAAAGCAAGACA 1048

541 CTACCTGTGATTTTATGATGACATGTGAGAAAAAGCAAGACA 586

RESULT 15

AAL13574

ID AAL13574 standard; cDNA; 783 BP.

XX

AC AAL13574;

XX 07-DEC-2001 (first entry)

XX Human breast cancer expressed polynucleotide 6031.

XX Human; breast cancer; cell marker; cytostatic; ss.

XX Homo sapiens.

XX WO200151628-A2.

XX 19-JUL-2001.

XX 10-JAN-2001; 2001WO-US000798.

XX 14-JAN-2000; 2000US-0176077P.

XX 14-MAR-2000; 2000US-0189167P.

XX 24-MAR-2000; 2000US-0192099P.

XX 29-MAR-2000; 2000US-0193480P.

XX 15-MAY-2000; 2000US-0205230P.

XX 09-JUN-2000; 2000US-0211315P.

XX 25-JUL-2000; 2000US-0220534P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2001-451856/48.

XX New peptide useful as a marker for the diagnosis of breast cancer.

XX Claim 1; Page 1083; 3695pp; English.

XX The invention relates to human breast cancer expressed polynucleotides (AAL07544-AAL26789) and methods of assessing whether a patient is afflicted with breast cancer by examining the correlation between the expression of certain markers and the cancerous state of breast cells.

XX The polynucleotides and encoded polypeptides are potential markers for detecting, diagnosing, monitoring, characterising treating and potentially preventing breast cancer. The polynucleotides and encoded polypeptides are also useful for isolating compounds with cytostatic activity

XX Sequence 783 BP; 228 A; 157 C; 141 G; 248 T; 0 U; 9 Other;

Query Match

Best Local Similarity 23.3%; Score 546.2; DB 4; Length 783;

Matches 585; Conservative 0; Mismatches 10; Indels 4; Gaps 3;

460 AAGTACTACCCCAATAAATCTTCTTATGCTCTGGAAGAGAAAGCAATGTAATCACT 519

50 AGGTACTACCCCAATAAATCTTCTTATGCTCTGGAAGAGAAAGCAATGTAATCACT 109

520 ATGATTTATTCACAAACCAAGCAATCTCCACAAATTTTAACTATCTGATCTCTTCT 579

Db 110 ATGATTATTGCACAAACACACAGAAATCTCCAACAATTTTAAAGTAATCTGATCCTCTTCT 169
Qy 580 TCGAGAAATTTGTTACTATAGTCTTTCCCTTATGATGCTTATTAAGTCTTACTAGTATTAATC 639
Db 170 TCGAGAAATTTGTTACTATAGTCTTTCCCTTATGATGCTTATTAAGTCTTACTAGTATTAATC 229
Qy 640 AAATTTCTATAAAATTTCTTACTTAAAGTCTTTAAGAACTGGGTTCTTCCCTTTGATGTTATTC 699
Db 230 AAATTTCTATAAAATTTCTTACTTAAAGTCTTTAAGAACTGGGTTCTTCCCTTTGATGTTATTC 289
Qy 700 ATGTTCAAGAAAGAAACACACTTACTCTTTTAGGACAACTTCTAGAAATCTATAGTAGT 759
Db 290 ATGTTCAAGAAAGAAACACACTTACTCTTTTAGGACAACTTCTAGAAATCTATAGTAGT 349
Qy 760 ATCAGGATATATTTTGGCTTTTAAATATATTTTGGTTATTTTGAATACAGACATTTGGCTCC 819
Db 350 ATCAGGATATATTTTGGCTTTTAAATATATTTTGGTTATTTTGAATACAGACATTTGGCTCC 409
Qy 820 AAATTTTCATCTTTGCAATAGTAGTACTTTTCACTAGAACTTCTCAACATTTGGGAAC 879
Db 410 AAATTTTCATCTTTGCAATAGTAGTACTTTTCACTAGAACTTCTCAACATTTGGGAAC 469
Qy 880 TTTGCAATATGAGCATCATATGTTAAGGCTGTATCATTTAATGCTATGATACATT 939
Db 470 TTTGCAATATGAGCATCATATGTTAAGGCTGTATCATTTAATGCTATGATACATT 529
Qy 940 GTTTTCTCCCTATGCCAAACAGGTGAAC--AAACGTAGTTGTTTTTACTGATACATAAT 997
Db 530 GTTTTCTCCCTATGCCAAACAGGTGAACCAACGTAGTTGTTTTTACTGATACATAAT 589
Qy 998 GTTGGCTACCTGTGATTTTATAGTAGCATGTCTAG-AAAAAGGCAAGCAAAATGGCC 1055
Db 590 GTTGGCT-CCCGGATTTTATAGTAGCATGTCTANAAAAAGGCAAGCAAAATGGGC 647

Search completed: April 25, 2004, 08:46:21
Job time : 627.343 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2004, 23:54:27 ; Search time 6177.51 Seconds
(without alignments)
16460.143 Million cell updates/sec

Title: US-10-051-835-25
Perfect score: 2346
Sequence: 1 gccagaggggaaaaaagag.....aaacatttgatagttacat 2346

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb.pa.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rtd.*
36: em.htg_mam.*
37: em.htg_vrt.*
38: em.sy.*
39: em.htgo_hum.*
40: em.htgo_mus.*
41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	2332.4	99.4	64196	9	AC110804	AC110804 Homo sapi
C 2	2304.4	98.2	2321	6	AX281638	AX281638 Sequence
C 3	1613	68.8	1632	9	HSN806550	BS38293 Homo sapi
C 4	1599.4	68.2	1626	6	BD270271	BD270271 Human nuc
C 5	1599.4	68.2	1626	6	AX035207	AX035207 Sequence
C 6	1599.4	68.2	1626	6	AX342061	AX342061 Sequence
C 7	1599.4	68.2	1626	6	AX342121	AX342121 Sequence
C 8	1599.4	68.2	1626	6	AX441124	AX441124 Sequence
C 9	582.8	24.8	586	6	AX351446	AX351446 Sequence
C 10	581.4	24.8	575	6	AR273526	AR273526 Sequence
C 11	581.4	24.8	575	6	AR277107	AR277107 Sequence
C 12	581.4	24.8	575	6	AX368559	AX368559 Sequence
C 13	568	24.2	586	6	AX341176	AX341176 Sequence
C 14	551.4	23.5	665	6	AR407382	AR407382 Sequence
C 15	401.6	17.1	2000	9	AB040375	AB040375 Homo sapi
C 16	323	13.8	323	6	AX333323	AX333323 Sequence
C 17	323	13.8	323	6	AX407813	AX407813 Sequence
C 18	322.2	13.7	324	6	AX341711	AX341711 Sequence
C 19	321.2	13.7	324	6	AX341282	AX341282 Sequence
C 20	321	13.7	321	6	AX322023	AX322023 Sequence
C 21	319.6	13.6	324	6	AX341552	AX341552 Sequence
C 22	273.2	11.6	277	6	AX341912	AX341912 Sequence
C 23	243.6	10.4	247	6	AX340305	AX340305 Sequence
C 24	242.6	10.3	247	6	AX341879	AX341879 Sequence
C 25	242.6	10.3	247	6	AX341931	AX341931 Sequence
C 26	242.4	10.3	244	6	AX351360	AX351360 Sequence
C 27	229.6	9.8	248	6	AX341175	AX341175 Sequence
C 28	109.8	4.7	204366	2	AC118181	AC118181 Rattus no
C 29	109.8	4.7	223893	2	AC106406	AC106406 Rattus no
C 30	93.4	4.0	196358	2	AC101992	AC101992 Mus muscu
C 31	69.6	3.0	67370	3	PFMALP3	AL031746 Plasmodiu
C 32	64	2.7	309	6	AX341780	AX341780 Sequence
C 33	63	2.7	1192	9	HSR323759	AX323759 Homo sapi
C 34	62.4	2.7	1141	6	AX083744	AX083744 Sequence
C 35	61.8	2.6	13138	3	AC116305_4	Continuation (5 of
C 36	61.8	2.6	110000	3	AC116305_3	Continuation (4 of
C 37	61.6	2.6	349751	3	PFMALP3	AL035476 Plasmodiu
C 38	61	2.6	8056	6	AX599046	AX599046 Sequence
C 39	60.4	2.6	349751	3	PFMALP3	AL035476 Plasmodiu
C 40	60.2	2.6	164900	2	BS571776	BS571776 Danio rer
C 41	60.2	2.6	349980	6	AX344559	AX344559 Sequence
C 42	59.8	2.5	89665	9	AC063976	AC063976 Homo sapi
C 43	59.8	2.5	163660	2	AC046165	AC046165 Homo sapi
C 44	59.6	2.5	250029	3	AE014839	AE014839 plasmodiu
C 45	59.2	2.5	700	6	AX183285	AX183285 Sequence

ALIGNMENTS

RESULT 1
AC110804/c
LOCUS AC110804 64196 bp DNA linear PRI 30-MAR-2002
DEFINITION Homo sapiens BAC clone RP11-725C19 from 4, complete sequence.
ACCESSION AC110804
VERSION AC110804.3 GI:19570222
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 64196)
AUTHORS Sulston,J.E. and Waterston,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)

```

99063792
9847074
2 (bases 1 to 64196)
PEARMAN, C., KOZLOVICZ, A. and DIGNAN, G.
The sequence of Homo sapiens BAC clone RP11-725C19
Unpublished (2001)
3 (bases 1 to 64196)
Waterston, R. H.
Direct Submission
Submitted (15-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 64196)
Waterston, R. H.
Direct Submission
Submitted (06-MAR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 64196)
Waterston, R. H.
Direct Submission
Submitted (21-MAR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 64196)
Waterston, R.
Direct Submission
Submitted (30-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Mar 21, 2002 this sequence version replaced gi:19073866.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
-----
Center project name: H_NH0725C19
-----

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPI1-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E.,
Tateno, M., Catanese, J. J. and de Jong, P. J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at http://www.chori.org
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-8P17, 2000 bp overlap.
Actual start of this clone is at base position 185268 of
RP11-393A12; actual end is at base position 64196 of RP11-725C19.

```

```

FEATURES             Location/Qualifiers
     source            1..64196
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /map="14"
                     /clone="RP11-725C19"
     repeat_region     1..39
                     /rpt_family="AT_rich"
     repeat_region     162..182
                     /rpt_family="AT_rich"
     repeat_region     312..607
                     /rpt_family="Alu"
     misc_feature       369..505
                     /note="match to EST BG257959 (NID:gi2767775)"
     repeat_region     877..909
                     /rpt_family="AT_rich"
     repeat_region     1113..1536
                     /rpt_family="L2"
     repeat_region     2400..2427
                     /rpt_family="AT_rich"
     repeat_region     2941..2984
                     /rpt_family="L2"
     repeat_region     3413..3435
                     /rpt_family="(TA)n"
     misc_feature       3829..4216
                     /note="match to EST H06135 (NID:g869687) Y177b12.s1"
     misc_feature       3829..4170
                     /note="match to EST BI497113 (NID:gi5336457)"
     misc_feature       3831..4157
                     /note="match to EST AA639698 (NID:g2563477) np60f09.s1"
     misc_feature       3835..4360
                     /note="match to EST AW237145 (NID:g6569534) xm66b03.x1"
     misc_feature       3836..4376
                     /note="similar to EST BM791320 (NID:gi19139552)"
     misc_feature       3837..4405
                     /note="similar to Bos taurus EST BI539797 (NID:gi5380907)"
     misc_feature       3844..4610
                     /note="match to EST BI857800 (NID:gi5998547)"
     misc_feature       3933..4394
                     /note="match to EST AI694383 (NID:gi4971723) wd83b04.x1"
     misc_feature       4105..4843
                     /note="match to EST BG177619 (NID:gi2684322)"
     misc_feature       4137..4527
                     /note="match to EST AI267688 (NID:g3886855) aq93c10.x1"
     misc_feature       4153..4433
                     /note="match to EST C00538 (NID:gi432768)"
     misc_feature       4233
                     /note="match to EST BI497113 (NID:gi5336457)"
     misc_feature       4359..4746
                     /note="match to EST AA227760 (NID:gi849333) zr56a07.r1"
     misc_feature       4404..5178
                     /note="match to EST BG291471 (NID:gi13049386)"
     misc_feature       4412..4619
                     /note="similar to Homo sapiens EST BF154977
                     (NID:gi1050160)"
     misc_feature       4657..5028
                     /note="match to EST R20307 (NID:gi774941) y935a09.r1"
     misc_feature       4727..5195
                     /note="similar to Bos taurus EST AW669076 (NID:g7525590)"
     misc_feature       4740..5803
                     /note="similar to EST BM547723 (NID:gi18781736)"
     misc_feature       4859..5265
                     /note="match to EST AW901105 (NID:g8065310)"
     misc_feature       5002..5687
                     /note="match to EST BG254117 (NID:gi2763933)"
     misc_feature       5004..5213
                     /note="match to EST BF084219 (NID:gi10878049)"
     misc_feature       5074..5447
                     /note="match to EST N80420 (NID:gi1243121) y292e06.r1"

```

Polymorphisms exist between AC013365 and AC110804.

misc_feature 5255..6002 /note="match to EST BI869486 (NID:g16043159)"
misc_feature 5442..6151 /note="match to EST AW961203 (NID:g8150887)"
misc_feature 5497..5830 /note="match to EST AI540897 (NID:g4458270)"
misc_feature 5782..6124 /note="match to EST AA309186 (NID:g1961511)"
misc_feature 5875..6701 /note="match to EST BI334972 (NID:g15019629)"
misc_feature 5995..6331 /note="match to EST AA471165 (NID:g2199022)"
misc_feature 6000..6124 /note="match to EST AW996896 (NID:g8257130)"
misc_feature 6032..6836 /note="similar to EST BM474943 (NID:g18523985)"
misc_feature 6140..6518 /note="match to EST AW996897 (NID:g8257131)"
misc_feature 6444..6885 /note="match to EST AA375957 (NID:g2028275)"
repeat_region 6878..7187 /rpt_family="Alu"
misc_feature 6899..7362 /note="match to EST AA488687 (NID:g2216118) ab38f03.s1"
misc_feature 6970..7746 /note="match to EST BE739933 (NID:g10153837)"
misc_feature 7099..7815 /note="match to EST BF381953 (NID:g11363256)"
misc_feature 7124..7125 /note="match to EST AW996897 (NID:g8257131)"
misc_feature 7174..7495 /note="match to EST BF297528 (NID:g12325656)"
misc_feature 7253..7884 /note="similar to Homo sapiens EST BG499179 (NID:g13460696)"
misc_feature 7568..7949 /note="match to EST AA307001 (NID:g1959330)"
misc_feature 7613..8332 /note="similar to EST BM706727 (NID:g19019985)"
repeat_region 7836..8011 /rpt_family="AchHob"

Query Match 99.4%; Score 2332.4; DB 9; Length 64196;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2344; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 GCCAGAGGGGAAAAAAGAGTAATGCACAGGTATCTCTTTTGCAGTGGTGATGATTTT 60
DB 6179 GCCAGAGGGGAAAAAAGAGTAATGCACAGGTATCTCTTTTGCAGTGGTGATGATTTT 6120

QY 61 GAGTACTTGTGTGACAGGTATTTATACAGCATCTCTGCGAAACCTATTAGGCCTT 120
DB 6119 GAGTACTTGTG-TGTGACAGGGTAATATTACAGCATCTTTGCGAAACCTATTAGGCCTT 6061

QY 121 TGCATGTTTAAAGCTGTATATTTTGGTGTGTGAGTGGTCTGACITTAATCTGTATTAT 180
DB 6060 TGCATGTTTAAAGCTGTATATTTTGGTGTGTGAGTGGTCTGACITTAATCTGTATTAT 6001

QY 181 AAAATTTAGACATCAAAATTTTCTACTACTAACTTTATAGATGATATCTTGGAGCAC 240
DB 6000 AAAATTTAGACATCAAAATTTTCTACTACTAACTTTATAGATGATATCTTGGAGCAC 5941

QY 241 AGTCATATCACATCGGAGGCAATGCAATGTGTGTACCTGGTCTGAGTTTGAACGTCT 300
DB 5940 AGTCATATCACATCGGAGGCAATGCAATGTGTGTACCTGGTCTGAGTTTGAACGTCT 5881

QY 301 TATTTCAAAGATTTCTGAATTAATTTTCCCTAGAAATTTCTCTTCATCTCCAAAGTACA 360
DB 5880 TATTTCAAAGATTTCTGAATTAATTTTCCCTAGAAATTTCTCTTCATCTCCAAAGTACA 5821

QY 361 AACATCTTTGAAGATGAACAGATTTGTTCCCATGAATGTATGTCTCACTCTGACTAGA 420
DB 5820 AACATCTTTGAAGATGAACAGATTTGTTCCCATGAATGTATGTCTCACTCTGACTAGA 5761

QY 421 AACGATCTATGTTAAATGACTGTGTATATGAATTTTCAAGTACTACCCCAATAACTT 480
DB 5760 AACGATCTATGTTAAATGACTGTGTATATGAATTTTCAAGTACTACCCCAATAACTT 5701

QY 481 TCTTATTGCTCTGAAAGAAAGAAAGCAATGTAATCACTATATTTTGCACAAACCAACC 540
DB 5700 TCTTATTGCTCTGAAAGAAAGAAAGCAATGTAATCACTATATTTTGCACAAACCAACC 5641

QY 541 AGAATTTCTCCAACTTTTAAAGTAACTGATCTCTCTTTGGAGAAATTTGTACCTAAT 600
DB 5640 AGAATTTCTCCAACTTTTAAAGTAACTGATCTCTCTTTGGAGAAATTTGTACCTAAT 5581

QY 601 AGTATTTCTCTTATGAATGTTTATTTACTACTGGTATAAATCAAAATTTCTATAAAATTTCTCTAC 660
DB 5580 AGTATTTCTCTTATGAATGTTTATTTACTACTGGTATAAATCAAAATTTCTATAAAATTTCTCTAC 5521

QY 661 TTAAGTCTTTAAGAACTGGGTTCTTCTTTGATGTTTATTTCAATGTTTCAAGAAAGAAACAACA 720
DB 5520 TTAAGTCTTTAAGAACTGGGTTCTTCTTTGATGTTTATTTCAATGTTTCAAGAAAGAAACAACA 5461

QY 721 CTTTACTCTTTTAAAGCAATTTCTAGAACTATAGTAGTATCAGGATATATTTTGTCTTA 780
DB 5460 CTTTACTCTTTTAAAGCAATTTCTAGAACTATAGTAGTATCAGGATATATTTTGTCTTA 5401

QY 781 AAATATATTTTGGTATTTTGAATACAGCAATTTGGCTCCAAATTTTCACTTTTGCACAAAT 840
DB 5400 AAATATATTTTGGTATTTTGAATACAGCAATTTGGCTCCAAATTTTCACTTTTGCACAAAT 5341

QY 841 AGTATGACTTTTCACTAGAACTTTCTCAACTTTGGGAACTTTGCAATATGACATCATATA 900
DB 5340 AGTATGACTTTTCACTAGAACTTTCTCAACTTTGGGAACTTTGCAATATGACATCATATA 5281

QY 901 TGTGTTAAGGCTGTATCATTTAATGCTATGAGATACATTTGTTTTCTCTCCCTATGCCAAACA 960
DB 5280 TGTGTTAAGGCTGTATCATTTAATGCTATGAGATACATTTGTTTTCTCTCCCTATGCCAAACA 5221

QY 961 GGTGAACAAACGCTAGTTGTTTTTACTGATCTAAATTTGGTCTACTGTGATTTTATAG 1020
DB 5220 GGTGAACAAACGCTAGTTGTTTTTACTGATCTAAATTTGGTCTACTGTGATTTTATAG 5161

QY 1021 TATGACATCTCAGAAAAGGCAAGCAAAATGCCCTTGTACTGTAATCTTCGGCAAC 1080
DB 5160 TATGACATCTCAGAAAAGGCAAGCAAAATGCCCTTGTACTGTAATCTTCGGCAAC 5101

QY 1081 TTATTTGGTCTTCAATTTTCTGACAGCAGGATTTGATCAATATTTGTAGAGCTTGGGTA 1140
DB 5100 TTATTTGGTCTTCAATTTTCTGACAGCAGGATTTGATCAATATTTGTAGAGCTTGGGTA 5041

QY 1141 GAATGGATTTACATGTTAGTATGACCTGGTAGAAATGGTTTTTGTATTGATCTCAGAAAT 1200
DB 5040 GAATGGATTTACATGTTAGTATGACCTGGTAGAAATGGTTTTTGTATTGATCTCAGAAAT 4981

QY 1201 TCATCTCAGGATGAACTCTTTTATGTCTTTTATGTATGTATGTATGTATGTATGTATGTAT 1260
DB 4980 TCATCTCAGGATGAACTCTTTTATGTCTTTTATGTATGTATGTATGTATGTATGTATGTAT 4921

QY 1261 AAGATGTTTTTAAAGCTTTTGTCTAAAAATTTGGCCTAGGAATGGTAACTTCAATTTTCA 1320
DB 4920 AAGATGTTTTTAAAGCTTTTGTCTAAAAATTTGGCCTAGGAATGGTAACTTCAATTTTCA 4861

QY 1321 GTTCCAAAGGGGTAGAAAAATAATATATGTGTGTGTATGTATGTATGTATGTATGTATGTAT 1380
DB 4860 GTTCCAAAGGGGTAGAAAAATAATATATGTGTGTGTATGTATGTATGTATGTATGTATGTAT 4801

QY 1381 GTACTCTATGAAATGTTTAAATATTTTTCATATTTCTGTGCAAGCAATTTTATAATTTTGT 1440
DB 4800 GTACTCTATGAAATGTTTAAATATTTTTCATATTTCTGTGCAAGCAATTTTATAATTTTGT 4741

QY 1441 CAACAAGTGGAGTCCATTTAGCCAGTGGGAAAGCTTTGGAACTCAGGTTACCCCTGAAG 1500
DB 4740 CAACAAGTGGAGTCCATTTAGCCAGTGGGAAAGCTTTGGAACTCAGGTTACCCCTGAAG 4681

1501 GATATGCTGGCAGCCATCTCTTTGATCTGCTTAACTGTAATTTATAGACCACTAAA 1560
1561 TCCTAACTTGGATCTGGATGCAATAGTATGACCTTGTACCATTTCCAGAAATTTCCAG 1620
1621 GGATCGTGGGTTTGGTCTAGTATGAAACACAAAGACAGAGATCCAGCTGAAAAA 1680
1681 GAGTGATCCTCAATATCCTAACTAACTGTCCTCACTCAAGCAGAGTTTCTCACTCTG 1740
1741 GCACTGTGATCATGAACCTTAGTAGAGGGATGCTGTATTTATPACAAATTTAATACA 1800
1801 ATGCTTTACATGATGAATCTTAAAGAGCAAACTGCAATTTATTTCTGATCCACAT 1860
1861 TCCAATCATATTAGAACTTAAGATATTTATCTATGAAGATATAAATGGTCCAGAGACTT 1920
1921 TCATCTGTGATGCGTTGTTCTTAGGTTCTAGCACTGATGCTGCTGCGGAGAGTTCTG 1980
1981 ATATGTAATATAAATGGATTTCTATAGCTAAATGAGTTCCCTCTGGGAGAGTTCTG 2040
2041 GTACTGCAATACAAATGCCAGATGGTGTATGAGGCTATTTGTAAGTAACTGTAAGA 2100
2101 TGCTATCAAGTAACTGATGTTGTTTTCATCTATGAAACTCTTGTATGATGCTGTTTG 2160
2161 TATGGAATAAATTTGGTGAATATGATGCTTCACTTGAATTTGAATTTGATGATG 2220
2221 TTGTAATTTATGATGATATACCTGTACGCTTCTAGTTGCTTCAACCATTTTATACCAT 2280
2281 TTTGTACATATTTTACTTGAAAAATTTTAAATGGAATTTAAATTAACATTTGATG 2340
2341 TTACAT 2346
2340 TTACAT 3835

AX281638 2321 bp DNA linear PAT 02-NOV-2001
AX281638 Sequence 47 from Patent WO0177389.
AX281638
AX281638.1 GI:16608889
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

Shiffman,D., Sonoyvi,R., Lawn,R., Seilhamer,J.J., Porter,G.J.,
Mikita,T. and Tai,J.
Genes expressed in foam cell differentiation
Patent: WO 0177389-A 47 18-OCT-2001;
Incyte Genomics, Inc. (US)
Location/Qualifiers
1..2321
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 200578.1"

Query Match 98.2%; Score 2304.4; DB 6; Length 2321;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2316; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 28 CAGTATCTCTTTGAGTGGTGTGACCTGATTTTGAAGCTGTTAAGCTGTATAATTTGTTG 87
Db 4 CAGTATCTCTTTGAGTGGTGTGACCTGATTTTGAAGCTGTTAAGCTGTATAATTTGTTG 62
QY 88 TACAGCATCTTTGGGAAAACCTATTAGGCTTTTGAAGCTGTTAAGCTGTATAATTTGTTG 147
Db 63 TACAGCATCTTTGGGAAAACCTATTAGGCTTTTGAAGCTGTTAAGCTGTATAATTTGTTG 122
QY 148 GGTGTGAGTGGTGTGATTTAAATGTTATTAATAATTTAGACATCAAAATTTTCTTACT 207
Db 123 GGTGTGAGTGGTGTGATTTAAATGTTATTAATAATTTAGACATCAAAATTTTCTTACT 182
QY 208 AACTAACTTTATAGATGTCATCTTGAAGCAGATCATATCATCTGGGAGCAATGCA 267
Db 183 AACTAACTTTATAGATGTCATCTTGAAGCAGATCATATCATCTGGGAGCAATGCA 242
QY 268 ATGTGGTTACCTGCTCCTAGTGTGAACTGTCTTTATTTCAAAAGATTTTGAATTAATTT 327
Db 243 ATGTGGTTACCTGCTCCTAGTGTGAACTGTCTTTATTTCAAAAGATTTTGAATTAATTT 302
QY 328 TTCCCTAGAAATTTCTCTTCAATTTCCAAAGTACAAACATATCTTTGAAGCAATGAAACAGATT 387
Db 303 TTCCCTAGAAATTTCTCTTCAATTTCCAAAGTACAAACATATCTTTGAAGCAATGAAACAGATT 362
QY 388 GTTCCCATGAATGTATGCTCATCTCGACTAGAAACGATCTATGTTAAATGACTGTGTAT 447
Db 363 GTTCCCATGAATGTATGCTCATCTCGACTAGAAACGATCTATGTTAAATGACTGTGTAT 422
QY 448 ATGAATATTTCAAGTACTACCCCAATTAATTTCTTTATTTGCTCTGAAGCAATGAAACAGCA 507
Db 423 ATGAATATTTCAAGTACTACCCCAATTAATTTCTTTATTTGCTCTGAAGCAATGAAACAGCA 482
QY 508 ATGTAATACATATGATTTATGCAAAACCAACCAAGATTTCTCCAACTTTTAACTAATC 567
Db 483 ATGTAATACATATGATTTATGCAAAACCAACCAAGATTTCTCCAACTTTTAACTAATC 542
QY 568 TGATCTCTCTTGGAGAAATTTGTTACCTTAATAGTTTCTTTCTTATGAATGTTATTAATA 627
Db 543 TGATCTCTCTTGGAGAAATTTGTTACCTTAATAGTTTCTTTCTTATGAATGTTATTAATA 602
QY 628 CTGGTATAAATCAAAATTTCTATAAATTTCTTACTTAAAGTCTTTAAGAACTGGGTTCTTCT 687
Db 603 CTGGTATAAATCAAAATTTCTATAAATTTCTTACTTAAAGTCTTTAAGAACTGGGTTCTTCT 662
QY 688 TTGATGTTATTCATGTTCTGAGAAAGAAACACATTTTACTCTTTTAGCAATTTCTTAGA 747
Db 663 TTGATGTTATTCATGTTCTGAGAAAGAAACACATTTTACTCTTTTAGCAATTTCTTAGA 722
QY 748 ATCTATAGTAGTATCAGGATATATTTTGTCTTTAAATATATTTTGGTATTTTGAATACA 807
Db 723 ATCTATAGTAGTATCAGGATATATTTTGTCTTTAAATATATTTTGGTATTTTGAATACA 782
QY 808 GACATTTGGCTCAAAATTTTCTATCTTTGCAAAATGATGACTTTTCTACTAGAACTTCTCA 867
Db 783 GACATTTGGCTCAAAATTTTCTATCTTTGCAAAATGATGACTTTTCTACTAGAACTTCTCA 842

```
QY 868 ACATTTGGGAACCTTTGCAAAATATGAGCATCATATGTGTAAAGGCTGTATCATTTAAAGCT 927
Db 843 ACATTTGGGAACCTTTGCAAAATATGAGCATCATATGTGTAAAGGCTGTATCATTTAAAGCT 902
QY 928 ATGAGATACATTTGTTTCTCCCTATGCGCAAAACAGGTGAACAAACGCTAGTCTGTTTCTTACT 987
Db 903 ATGAGATACATTTGTTTCTCCCTATGCGCAAAACAGGTGAACAAACGCTAGTCTGTTTCTTACT 962
QY 988 GATACATAATGTTGGCTACCTGTGTATTTTATAGTATGCAATGTCAGAAAAGGCAAGAC 1047
Db 963 GATACATAATGTTGGCTACCTGTGTATTTTATAGTATGCAATGTCAGAAAAGGCAAGAC 1022
QY 1048 AAATGGCCCTCTTGTAAGTCAATTTCTGCGCAAAACCTTATTTGGGCTCTTCAATTTCTGACAGAC 1107
Db 1023 AAATGGCCCTCTTGTAAGTCAATTTCTGCGCAAAACCTTATTTGGGCTCTTCAATTTCTGACAGAC 1082
QY 1108 AGGATTTGACATCAATTTTATAGTATGCGGTAGAAATGATATACATGCTAGTATGACACT 1167
Db 1083 AGGATTTGACATCAATTTTATAGTATGCGGTAGAAATGATATACATGCTAGTATGACACT 1142
QY 1168 GGTAGAAATGGTTTTAGTTATTTAGTCTCAGAAATTCATCTCAGGATGAATCTTTTATGTCT 1227
Db 1143 GGTAGAAATGGTTTTAGTTATTTAGTCTCAGAAATTCATCTCAGGATGAATCTTTTATGTCT 1202
QY 1228 TTTTATGTAAGCATATCTGAATTTACTTTTAAAGATGGTTTTAGAAAGCTTTTGCTTAA 1287
Db 1203 TTTTATGTAAGCATATCTGAATTTACTTTTAAAGATGGTTTTAGAAAGCTTTTGCTTAA 1262
QY 1288 AAATTTGGCCCTAGGAATGTAATCTTCAATTTTCAGTTGCGCAAGGGTACAAAATAATATG 1347
Db 1263 AAATTTGGCCCTAGGAATGTAATCTTCAATTTTCAGTTGCGCAAGGGTACAAAATAATATG 1322
QY 1348 TGTGTTGTTATGTTTATGTTAAATTTATTTATTTAGTATCTATCTATGAATTTTAAATAT 1407
Db 1323 TGTGTTGTTATGTTTATGTTAAATTTATTTATTTAGTATCTATCTATGAATTTTAAATAT 1382
QY 1408 TTTTCATATCTGTCAGCAAGCAATTTTAAATTTGCAACAGTGGAGTCCATTTAGCCCACT 1467
Db 1383 TTTTCATATCTGTCAGCAAGCAATTTTAAATTTGCAACAGTGGAGTCCATTTAGCCCACT 1442
QY 1468 GCGAAAGCTCTTGGAACTCAGGTTTACCTTTGAAGGATATGCTGGCAAGCATCTCTTTGATC 1527
Db 1443 GCGAAAGCTCTTGGAACTCAGGTTTACCTTTGAAGGATATGCTGGCAAGCATCTCTTTGATC 1502
QY 1528 TGTGCTTAACTGTAATTTATAGACAGTAAATCCCTAACTTGATCTGGAATGCAATTA 1587
Db 1503 TGTGCTTAACTGTAATTTATAGACAGTAAATCCCTAACTTGATCTGGAATGCAATTA 1562
QY 1588 GTTATGACCTTGTAACATTTCCAGAAATTTTCAAGGATCGTGGGTTTGGTCTAGTATG 1647
Db 1563 GTTATGACCTTGTAACATTTCCAGAAATTTTCAAGGATCGTGGGTTTGGTCTAGTATG 1622
QY 1648 AAAACACAAAGACAGAGATCCAGTGAAGAGTATCTCTCAATATCTTAACTAACT 1707
Db 1623 AAAACACAAAGACAGAGATCCAGTGAAGAGTATCTCTCAATATCTTAACTAACT 1682
QY 1708 GGTCTCTCACTCAAGCAGAGTTTCTTCACTCTGGCACTGTGATCATGAACTTAGTAGAG 1767
Db 1683 GGTCTCTCACTCAAGCAGAGTTTCTTCACTCTGGCACTGTGATCATGAACTTAGTAGAG 1742
QY 1768 GGGATTTGTGTATTTTATACAAATTTAATAAATGTTCTTCAATTTGATTAATTTCTTAAA 1827
Db 1743 GGGATTTGTGTATTTTATACAAATTTAATAAATGTTCTTCAATTTGATTAATTTCTTAAA 1802
QY 1828 GAGCAAACTGCAATTTTATTTCTGATCCACATTTCCAACTCATATTAGAACTAAGATATTT 1887
Db 1803 GAGCAAACTGCAATTTTATTTCTGATCCACATTTCCAACTCATATTAGAACTAAGATATTT 1862
QY 1888 ATCTATGAAGATATAAATGGTGCAGAGACATTTTCACTGTGGATTTGGTGTCTTCTTAG 1947
Db 1863 ATCTATGAAGATATAAATGGTGCAGAGACATTTTCACTGTGGATTTGGTGTCTTCTTAG 1922
QY 1948 GGTCTCTAGCACTGATCCCTGCAAGCATGTGATATGTGAATAAATGGATTTCTTCTA 2007
```

```
Db 1923 GGTCTCTAGCACTGATGCTCTGCAAGCATGTGATATGTGAATAAATGGATTTCTTCTA 1982
QY 2008 TAGCTAAATGAGTTCCCTCTGCGGAGAGTTCTGGTACTGCAATCAATGCCAGATGGTG 2067
Db 1983 TAGCTAAATGAGTTCCCTCTGCGGAGAGTTCTGGTACTGCAATCAATGCCAGATGGTG 2042
QY 2068 TTTATGGGCTATTGTTGTAAGTAAAGTGGTAAGATGCTATGAAGTAAAGTGTGTTGTTTTC 2127
Db 2043 TTTATGGGCTATTGTTGTAAGTAAAGTGGTAAGATGCTATGAAGTAAAGTGTGTTGTTTTC 2102
QY 2128 ATCTTAGGAACTCTTGATGATGCTTTTGTATGGAATAAATTTGGTGCATATGA 2187
Db 2103 ATCTTAGGAACTCTTGATGATGCTTTTGTATGGAATAAATTTGGTGCATATGA 2162
QY 2188 TGTCAATCAACTTTGATGATGCTTTTGTATGGAATTTTGTATGATATATACCTGTCA 2247
Db 2163 TGTCAATCAACTTTGATGATGCTTTTGTATGGAATTTTGTATGATATATACCTGTCA 2222
QY 2248 CGTCTTAGTGTCTCAACCACTTTATTAACCACTTTTGTATGATATATACCTGTCA 2307
Db 2223 CGTCTTAGTGTCTCAACCACTTTTATAACCACTTTTGTATGATATATATACCTGTCA 2282
QY 2308 TTTAAATGGAATTTTAAATAAACAATTTGATGATTTACAT 2346
Db 2283 TTTAAATGGAATTTTAAATAAACAATTTGATGATTTACAT 2321

RESULT 3
HSM806550
LOCUS Homo sapiens mRNA; cDNA DKF2p686F09157 (from clone DKF2p686F09157).
DEFINITION BX538293
VERSION BX538293.1 GI:31874775
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1632)
AUTHORS Lauber, J., Bahr, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fobo, G., Han, M. and Wiemann, S.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKF2p686F09157) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.

FEATURES
Location/Qualifiers
1..1632
/orxganism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="4Q26.3"
/clone="DKF2p686F09157"
/tissue_type="human colon endothel primary cell culture"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev stage="adult"
polyA_signal 1591..1596
polyA_site 1609
ORIGIN
Query Match 68.8%; Score 1613; DB 9; Length 1632;
Best Local Similarity 100.0%; Pred. No. 5.9e-279;
Matches 1613; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

734 GGCAATTCCTAGAACTCTATAGTAGTATCAGGATATATTTTGGTTTAAATATATTTTGG 793
1 GGCAATTCCTAGAACTCTATAGTAGTATCAGGATATATTTTGGTTTAAATATATTTTGG 60
794 TTAATTTGAATACAGACATTTGGCTCCAAATTTTTCATCTTTTGCAATAGTATGACTTTTC 853
61 TTAATTTGAATACAGACATTTGGCTCCAAATTTTTCATCTTTTGCAATAGTATGACTTTTC 120
854 ACTAGAACTCTCAACATTTGGGAATTTGCAATATGAGCATCATATGTTTAAAGCTG 913
121 ACTAGAACTCTCAACATTTGGGAATTTGCAATATGAGCATCATATGTTTAAAGCTG 180
914 TATCAATTAATGCTATGAGATACATTTGTTTCTCCCTATGCCAAACAGGTGAACAAAGCT 973
181 TATCAATTAATGCTATGAGATACATTTGTTTCTCCCTATGCCAAACAGGTGAACAAAGCT 240
974 AGTTGTTTTTACTGATCTAATATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1033
241 AGTTGTTTTTACTGATCTAATATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
1034 GAAAAAGGCAAGCAAAATGGCTCTTGTACTGATATCTTGGCAAACTTATTTGGGCTTTC 1093
301 GAAAAAGGCAAGCAAAATGGCTCTTGTACTGATATCTTGGCAAACTTATTTGGGCTTTC 360
1094 ATTTCTGACAGACAGATTTGATCTCAATATTTGTTAGAGCTTGGTGAATGATATACAT 1153
361 ATTTCTGACAGACAGATTTGATCTCAATATTTGTTAGAGCTTGGTGAATGATATACAT 420
1154 GGTAGTATGATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1213
421 GGTAGTATGATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 480
1214 AATCTTTTATGCTTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1273
481 AATCTTTTATGCTTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 540
1274 AAAGCTTTGCTAAATTTGGCTAGGAATGTTTCTTCACTTTCAGTTGCCAGGGGT 1333
541 AAAGCTTTGCTAAATTTGGCTAGGAATGTTTCTTCACTTTCAGTTGCCAGGGGT 600
1334 AGAAAAATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1393
601 AGAAAAATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 660
1394 ATGTATTTAAATATTTTCTATATCTGTCAGAGCATTTTATATTTGCAACAGTGGAGT 1453
661 ATGTATTTAAATATTTTCTATATCTGTCAGAGCATTTTATATTTGCAACAGTGGAGT 720
1454 CCATTAGCCAGTGGGAAAGCTTGGAACTCAGGTTACCTTTGAGGATATGCTGGCAG 1513
721 CCATTAGCCAGTGGGAAAGCTTGGAACTCAGGTTACCTTTGAGGATATGCTGGCAG 780
1514 CCATCTTTGATCTGCTTAACTGTTAATTTATAGCCAGTAAATCCCTTAACTTGA 1573
781 CCATCTTTGATCTGCTTAACTGTTAATTTATAGCCAGTAAATCCCTTAACTTGA 840
1574 TCTGGAATGATTTAGTTATGATCTTACCATTTCCAGAAATTTCCGGGATCGTGGGTT 1633
841 TCTGGAATGATTTAGTTATGATCTTACCATTTCCAGAAATTTCCGGGATCGTGGGTT 900
1634 TGCTCTAGTATTTGAAACACAGACAGAGATCCAGCTGAAAGAGTATCTCTCA 1693
901 TGCTCTAGTATTTGAAACACAGACAGAGATCCAGCTGAAAGAGTATCTCTCA 960
1694 TATCTTAACTTAACTGCTTCACTCAAGCAGAGTTTCTTCACTCTGGCAGTGTGATCAT 1753
961 TATCTTAACTTAACTGCTTCACTCAAGCAGAGTTTCTTCACTCTGGCAGTGTGATCAT 1020
1754 GAACTTAGTAGGGGATTTGTTGTTTATACAAATTTAATACATGCTTCAATG 1813
1021 GAACTTAGTAGGGGATTTGTTGTTTATACAAATTTAATACATGCTTCAATG 1080
1814 ATAAATTTCTTAAAGAGCAAACTGCAATTTTATTTCTGTCATCCCAATTCATATTA 1873

Db 1081 ATAAAAATTTCTTAAAGACAAACATGCAATTTATTTCTGCATCCACATTCCAATCATATTA 1140
Qy 1874 GAACTAAGATATTTATCTATCAAGATATAAATGTTGTCAGAGAGACTTTTCATCTGTGATTT 1933
Db 1141 GAACTAAGATATTTATCTATGAAGATATAAATGTTGTCAGAGAGACTTTTCATCTGTGATTT 1200
Qy 1934 GCGTTGTTTCTTAGGGTTCCTAGCAGTATGCTGCAACAGCATGTCATATGTGAATATA 1993
Db 1201 GCGTTGTTTCTTAGGGTTCCTAGCAGTATGCTGCAACAGCATGTCATATGTGAATATA 1260
Qy 1994 AATGGATTTCTTATAGTATAAATGAGTTCCCTCTGCGGAGAGTTCTGCTACTGCAATCAC 2053
Db 1261 AATGGATTTCTTATAGTATAAATGAGTTCCCTCTGCGGAGAGTTCTGCTACTGCAATCAC 1320
Qy 2054 AATGCCAGATGGTGTATGCGCTATTTGTTGTTAGTAAAGTGGTAAGATGCTATGAAGTAA 2113
Db 1321 AATGCCAGATGGTGTATGCGCTATTTGTTGTTAGTAAAGTGGTAAGATGCTATGAAGTAA 1380
Qy 2114 GTGTGTTTGTGTTTTCATCTTATGAAACTCTTGATGATGTCGTTTGTATGGAATAAAT 2173
Db 1381 GTGTGTTTGTGTTTTCATCTTATGAAACTCTTGATGATGTCGTTTGTATGGAATAAAT 1440
Qy 2174 TTGTGCAATATGATGTCATTCACCTTTGCAATTTGCAATTTGTTGTTGTTTATATG 2233
Db 1441 TTGTGCAATATGATGTCATTCACCTTTGCAATTTGCAATTTGTTGTTGTTTATATG 1500
Qy 2234 TATTATACCTGTCACGCTTCTAGTTGCTTCAACCATTTTAAACCATTTTGTATGACATAT 2293
Db 1501 TATTATACCTGTCACGCTTCTAGTTGCTTCAACCATTTTAAACCATTTTGTATGACATAT 1560
Qy 2294 TTACTGAAATATTTTAAATGGAATTTAAATTAACCATTTGATAGTTTACAT 2346
Db 1561 TTACTGAAATATTTTAAATGGAATTTAAATTAACCATTTGATAGTTTACAT 1613

RESULT 4
BD270271 Human nucleic acid sequence and protein sequence derived from
LOCUS BD270271 1626 bp DNA linear PAT 17-JUL-2003
DEFINITION Human nucleic acid sequence and protein sequence derived from
endothelial cell.
ACCESSION BD270271
VERSION BD270271.1 GI:33080039
KEYWORDS JP 2002537832-A/24.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1626)
AUTHORS Thierach, K.H., Glienke, J., Hinzmann, B. and Pilarsky, C.
TITLE Human nucleic acid sequence and protein sequence derived from
endothelial cell
JOURNAL Patent: JP 2002537832-A 24 12-NOV-2002;
SCHERING AKTIEGESELLSCHAFT
COMMENT OS Homo sapiens (human)
PN JP 2002537832-A/24
PD 12-NOV-2002
PF 08-MAR-2000 JP 2000603357
PR 08-MAR-1999 DE 199 11 684.9.01-OCT-1999 DE 199 48 679.4 PI
KARL HEINZ THIERAUCH, JENS GLIENKE, BERND HINZMANN, CHRISTIAN PI
PILARSKY
PC C12N15/09, A61K31/711, A61K38/00, A61K48/00, A61P1/16, A61P3/10, PC
A61P9/10,
PC A61P9/10, A61P13/12, A61P17/06, A61P19/02, A61P25/00, PC
A61P27/02,
PC A61P27/06, A61P29/00, A61P35/00, A61P41/00, A61P43/00, A61P43/00,
PC C07K14/47,
PC C07K16/18, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12P21/08, PC
C12N15/00,
PC C12N5/00, A61K37/02
CC Human nucleic acid sequence and protein sequence derived from
endothelial
CC cell

ORIGIN

Query Match 68.2%; Score 1599.4; DB 6; Length 1626;
Best Local Similarity 99.9%; Pred. No. 1.6e-276;
Matches 1611; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 734 GGCAATTCCTAGAACTCTAGTAGTATCAGGATATATTTTGGTAAATATATTTGG 793
DB 1 GGCAATTCCTAGAACTCTAGTAGTATCAGGATATATTTTGGTAAATATATTTGG 60
QY 794 TTATTTTGAATACAGACATTTGGCTCCAAATTTTTCATCTTTTGGCAATAGTAGCTTTTC 853
DB 61 TTATTTTGAATACAGACATTTGGCTCCAAATTTTTCATCTTTTGGCAATAGTAGCTTTTC 120
QY 854 ACTAGAACTTCTCAACATTTTGGCACTTTTGCATATAGAGCATCATCTGTGTTAGGCTG 913
DB 121 ACTAGAACTTCTCAACATTTTGGCACTTTTGCATATAGAGCATCATCTGTGTTAGGCTG 180
QY 914 TATCATTTAATAGTATGAGATACATTTGTTTCTCCCTATGCCAAACAGGTGAACAAACGT 973
DB 181 TATCATTTAATAGTATGAGATACATTTGTTTCTCCCTATGCCAAACAGGTGAACAAACGT 240
QY 974 AGTTGTTTTTCTGATACATTAATTTTGGCTACCTGTGATTTTATAGTATGACATGTCA 1033
DB 241 AGTTGTTTTTCTGATACATTAATTTTGGCTACCTGTGATTTTATAGTATGACATGTCA 300
QY 1034 GAAAAAGGCAAGCAAAATGGCTCTTGTACTGAATACTTCCGCAAACTTTATGGGCTTC 1093
DB 301 GAAAAAGGCAAGCAAAATGGCTCTTGTACTGAATACTTCCGCAAACTTTATGGGCTTC 360
QY 1094 ATTTTCTGACAGACAGATTTGACTCAATATTTGTAGAGCTTCCGAGAACTGATACAT 1153
DB 361 ATTTTCTGACAGACAGATTTGACTCAATATTTGTAGAGCTTCCGAGAACTGATACAT 420
QY 1154 GGTAGTACACTGGTAGAAATGGTTTTTATTTAGTATGATCAGAACTTCTCAGGATG 1213
DB 421 GGTAGTACACTGGTAGAAATGGTTTTTATTTAGTATGATCAGAACTTCTCAGGATG 480
QY 1214 AATCTTTTATGCTTTTATTTAGTATGATCAGAACTTCTCAGTATGATGATGATTTAG 1273
DB 481 AATCTTTTATGCTTTTATTTAGTATGATCAGAACTTCTCAGTATGATGATGATTTAG 540
QY 1274 AAAGCTTTTGTCTAAAAATTTGGCTAGGAATGTTAACTTCAATTTTCCAGTTGCCAAGGGT 1333
DB 541 AAAGCTTTTGTCTAAAAATTTGGCTAGGAATGTTAACTTCAATTTTCCAGTTGCCAAGGGT 600
QY 1334 AGAAAAATATATGTTGTTTATTTAGTATGATCAGAACTTCTCAGTATGATGATGATTTAG 1393
DB 601 AGAAAAATATATGTTGTTTATTTAGTATGATCAGAACTTCTCAGTATGATGATGATTTAG 660
QY 1394 ATGTATTTAATATTTTCTCATATTTCTGTGACAGCAATTTATATTTTGCACAACTGGAGT 1453
DB 661 ATGTATTTAATATTTTCTCATATTTCTGTGACAGCAATTTATATTTTGCACAACTGGAGT 720
QY 1454 CCATTTAGCCAGTGGGAAAGTCTTGGAACTCAGTTTACCTTTGAAGGATATGCTGGCAG 1513
DB 721 CCATTTAGCCAGTGGGAAAGTCTTGGAACTCAGTTTACCTTTGAAGGATATGCTGGCAG 780
QY 1514 CCATCTCTTTGATCTGTGCTTAACTGTAATTTATAGCAGCAATTAATCCCTAACTTTGA 1573
DB 781 CCATCTCTTTGATCTGTGCTTAACTGTAATTTATAGCAGCAATTAATCCCTAACTTTGA 840
QY 1574 TCTGGAATGATATGATTTATGACCTTTGACATTTCCAGAAATTTCCAGGGATCTGGGTT 1633
DB 841 TCTGGAATGATATGATTTATG-CCTTTGATCCATTTCCAGAAATTTCCAGGGATCTGGGTT 899
QY 1634 TGGTCTAGTATGATTAACACAGACAGAGATCCAGCTGAAAAAGAGTATCTCTCAA 1693
DB 900 TGGTCTAGTATGATTAACACAGACAGAGATCCAGCTGAAAAAGAGTATCTCTCAA 959
QY 1694 TATCCTAACTAACTGGTCTCTCAACTCAAGCAGAGTTTCTTCACTCTGGCACTGTGATCAT 1753
DB 960 TATCCTAACTAACTGGTCTCTCAACTCAAGCAGAGTTTCTTCACTCTGGCACTGTGATCAT 1019

QY 1754 GAACTTTAGTAGAGGGGATTTGTGTATTTTATACAAATTTTATACAACTCTTACATTTG 1813
DB 1020 GAACTTTAGTAGAGGGGATTTGTGTATTTTATACAAATTTTATACAACTCTTACATTTG 1079
QY 1814 ATAAAAATTTCTTAAAGAGCAAAACTGCAATTTTATTTCTGCAATCCACATTCOAATCATATTA 1873
DB 1080 ATAAAAATTTCTTAAAGAGCAAAACTGCAATTTTATTTCTGCAATCCACATTCOAATCATATTA 1139
QY 1874 GAACATAGATATTTATCTATCAAGATATAAATGGTGCAGAGAGACTTTCATCTGTGGATT 1933
DB 1140 GAACATAGATATTTATCTATCAAGATATAAATGGTGCAGAGAGACTTTCATCTGTGGATT 1199
QY 1934 GCGTTGTTTCTTAGGGTTCCTAGCACTGATGCCCTGCACAAGCATGTGATATGTGAATAA 1993
DB 1200 GCGTTGTTTCTTAGGGTTCCTAGCACTGATGCCCTGCACAAGCATGTGATATGTGAATAA 1259
QY 1994 AATGGATCTTCTATAGCTAAATGAGTTCCCTCTGGGAGAGTTCTGGTACTCCATCAC 2053
DB 1260 AATGGATCTTCTATAGCTAAATGAGTTCCCTCTGGGAGAGTTCTGGTACTCCATCAC 1319
QY 2054 AATGCCAGATGGTGTATGGGCTATTTGTGAATGTAAGTGAAGTGAAGTGAAGTGAAGTGA 2113
DB 1320 AATGCCAGATGGTGTATGGGCTATTTGTGAATGTAAGTGAAGTGAAGTGAAGTGAAGTGA 1379
QY 2114 GTGTGTTTGTTCATCTTATGGAACCTCTGATGCAATGTCTTTGTATGGAATAAAT 2173
DB 1380 GTGTGTTTGTTCATCTTATGGAACCTCTGATGCAATGTCTTTGTATGGAATAAAT 1439
QY 2174 TTGTGCAATATGATGTCAATCACTTTGCAATTTGAATTTGTTGTTGTTGTTGTTGTTGTTG 2233
DB 1440 TTGTGCAATATGATGTCAATCACTTTGCAATTTGAATTTGTTGTTGTTGTTGTTGTTGTTG 1499
QY 2234 TATTTATACCTGTGACGGTCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2293
DB 1500 TATTTATACCTGTGACGGTCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1559
QY 2294 TTACTTGAATAATTTTAAATGGAATTTAAATTAACATTTGATGTTTACAT 2346
DB 1560 TTACTTGAATAATTTTAAATGGAATTTAAATTAACATTTGATGTTTACAT 1612

RESULT 6

AX342061 1626 bp DNA linear PAT 12-JAN-2002
Sequence 24 from Patent EP1166798.

DEFINITION AX342061

ACCESSION AX342061

VERSION AX342061.1 GI:18151630

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Siemeister, G., Haberey, M. and Thierach, K.H.

TITLE Combinations and compositions which interfere with vegf/vegfr and

angiopoietin/tie receptor function and their use

JOURNAL Patent: EP 1166798-A 24 02-JAN-2002;

Schering Aktiengesellschaft (DE)

Location/Qualifiers

1. 1626

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 68.2%; Score 1599.4; DB 6; Length 1626;

Best Local Similarity 99.9%; Pred. No. 1.6e-276;

Matches 1611; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 734 GGCAATTCCTAGAACTCTAGTAGTATCAGGATATATTTTGGTAAATATATTTGG 793

DB 1 GGCAATTCCTAGAACTCTAGTAGTATCAGGATATATTTTGGTAAATATATTTGG 60

QY 794 TTATTTTGAATPACAGACATGGCTCCCAAAATTTTTCATCTTTTGCAAAATAGTATGACTTTTC 853
Db 61 TTATTTTGAATPACAGACATGGCTCCCAAAATTTTTCATCTTTTGCAAAATAGTATGACTTTTC 120
QY 854 ACTAGAACTCTTCAAATTTGGGAACTTTGCAAAATAGTATGACTTTTC 913
Db 121 ACTAGAACTCTTCAAATTTGGGAACTTTGCAAAATAGTATGACTTTTC 180
QY 914 TATCATTTAATGCTATGAGATACATCTTTTCTCCCTATGCGCAACAGGTGAAACAAACGT 973
Db 181 TATCATTTAATGCTATGAGATACATCTTTTCTCCCTATGCGCAACAGGTGAAACAAACGT 240
QY 974 AGTTGTTTTTACGATACATAAATGTTGGCTACCTGTGATTTTATAGTATGCAATGTCA 1033
Db 241 AGTTGTTTTTACGATACATAAATGTTGGCTACCTGTGATTTTATAGTATGCAATGTCA 300
QY 1034 GAAAAAGCAAGCAAAATGGGCTCTGCTACTGAAATCTTGGCAAACTTATGGGCTTC 1093
Db 301 GAAAAAGCAAGCAAAATGGGCTCTGCTACTGAAATCTTGGCAAACTTATGGGCTTC 360
QY 1094 ATTTTCTGACAGACAGGATTTGATCAATATTTGTAGAGCTTGGTGAATGATATCAT 1153
Db 361 ATTTTCTGACAGACAGGATTTGATCAATATTTGTAGAGCTTGGTGAATGATATCAT 420
QY 1154 GGTAGTATGACACTGGTAGAAATGGTTTGTATTTAGTATTTGACTCGAAATTCATCTCAGGATG 1213
Db 421 GGTAGTATGACACTGGTAGAAATGGTTTGTATTTAGTATTTGACTCGAAATTCATCTCAGGATG 480
QY 1214 AATCTTTTATGCTTTTATTTGTAAGCATATCTGAAATTTACTTTTATAAAGATGTTTTAG 1273
Db 481 AATCTTTTATGCTTTTATTTGTAAGCATATCTGAAATTTACTTTTATAAAGATGTTTTAG 540
QY 1274 AAAGCTTTGCTAAAAATTTGGGCTAGGAATGGTAACTTCAATTTTCACTTTGCGCAAGGGT 1333
Db 541 AAAGCTTTGCTAAAAATTTGGGCTAGGAATGGTAACTTCAATTTTCACTTTGCGCAAGGGT 600
QY 1334 AGAAAAATAATATGCTGTTGTTATTTGTTATTTGTTAAATATTTAGTACTATCTATGA 1393
Db 601 AGAAAAATAATATGCTGTTGTTATTTGTTATTTGTTAAATATTTAGTACTATCTATGA 660
QY 1394 ATGATTTTAAATATTTTCAATATTTCTGACAGCAATTTATTTTCAACAGTGGAGT 1453
Db 661 ATGATTTTAAATATTTTCAATATTTCTGACAGCAATTTATTTTCAACAGTGGAGT 720
QY 1454 CCATTTAGCCAGTGGGAAAGTCTTGGAACTCAGGTTACCCCTTCAAGAGATATGCTGGCAG 1513
Db 721 CCATTTAGCCAGTGGGAAAGTCTTGGAACTCAGGTTACCCCTTCAAGAGATATGCTGGCAG 780
QY 1514 CCACTCTTTGATCTGCTTAAACTGTAATTTATAGACAGCTAAATCCCTTAACTTTGGA 1573
Db 781 CCACTCTTTGATCTGCTTAAACTGTAATTTATAGACAGCTAAATCCCTTAACTTTGGA 840
QY 1574 TCTGGAATGCAATTTAGTATGACCTTTGATACCTTCCAGAAATTTCCAGGGGCAATCGTGGTT 1633
Db 841 TCTGGAATGCAATTTAGTATGACCTTTGATACCTTCCAGAAATTTCCAGGGGCAATCGTGGTT 899
QY 1634 TGGTCTAGTATGTAAGAAACACAGAAACAGAGATCCAGCTGAAAAAGAGATGATCTCAA 1693
Db 900 TGGTCTAGTATGTAAGAAACACAGAAACAGAGATCCAGCTGAAAAAGAGATGATCTCAA 959
QY 1694 TATCTTAATGCTGCTCTCACTCAGACAGAGTTTCTTCACTCTGSCATGATCAT 1753
Db 960 TATCTTAATGCTGCTCTCACTCAGACAGAGTTTCTTCACTCTGSCATGATCAT 1019
QY 1754 GAAACTTAGTAGAGGGGATTTGTGTATTTTATACAAATTTTAAATGCTTTTACATG 1813
Db 1020 GAAACTTAGTAGAGGGGATTTGTGTATTTTATACAAATTTTAAATGCTTTTACATG 1079
QY 1814 ATAAATTTCTTAAGAGCAAACTGCAATTTTCTGCTCCCAATTTCCAAATCATATTA 1873
Db 1080 ATAAATTTCTTAAGAGCAAACTGCAATTTTCTGCTCCCAATTTCCAAATCATATTA 1139

QY 1874 GAACTAAGATATTTATCTATGAGATATAAATGTTGCGAGAGACTTTTCATCTGTGGATT 1933
Db 1140 GAACTAAGATATTTATCTATGAGATATAAATGTTGCGAGAGACTTTTCATCTGTGGATT 1199
QY 1934 GGGTTGTTTTCTAGGGTTCCCTAGCACTGATCCCTGCAAGCATGTGATATGTAATAA 1993
Db 1200 GGGTTGTTTTCTAGGGTTCCCTAGCACTGATCCCTGCAAGCATGTGATATGTAATAA 1259
QY 1994 AATGATTTCTTCTATGCTAATAATGAGTTCCCTTGGGGAGAGTTCTGTACTGCAATCAC 2053
Db 1260 AATGATTTCTTCTATGCTAATAATGAGTTCCCTTGGGGAGAGTTCTGTACTGCAATCAC 1319
QY 2054 AATGCCAGATGGTGTATGAGTATTTGCTTATGCTTGAAGTAAAGTCTATGAAGTAA 2113
Db 1320 AATGCCAGATGGTGTATGAGTATTTGCTTATGCTTGAAGTAAAGTCTATGAAGTAA 1379
QY 2114 GTGCTGTTTTCTTCTATGCTTATGGAACCTCTTGTATGCTATGCTTGTATGGAATAAAT 2173
Db 1380 GTGCTGTTTTCTTCTATGCTTATGGAACCTCTTGTATGCTATGCTTGTATGGAATAAAT 1439
QY 2174 TTGGTGCATATGATGCTCATTCACCTTTGCTTGAATTTGATTTTATATG 2233
Db 1440 TTGGTGCATATGATGCTCATTCACCTTTGCTTGAATTTGATTTTATATG 1499
QY 2234 TATTATACCTGTCAGCTTCTAGTTGCTTCAACCAATTTTATAACCAATTTTGTACATAT 2293
Db 1500 TATTATACCTGTCAGCTTCTAGTTGCTTCAACCAATTTTATAACCAATTTTGTACATAT 1559
QY 2294 TTACTTGAATAATTTTAAATGGAATTTAAATGAACATTTTGAATGTTTACAT 2346
Db 1560 TTACTTGAATAATTTTAAATGGAATTTAAATGAACATTTTGAATGTTTACAT 1612

RESULT 7
AX342121 1626 bp DNA linear PAT 12-JAN-2002
LOCUS AX342121
DEFINITION Sequence 24 from Patent EP1166799.
ACCESSION AX342121
VERSION AX342121.1 GI:18151689
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Siemeister, G., Haberey, M. and Thierauch, K.H.
TITLE Combinations and compositions which interfere with vegf/vegf and
angiopoietin/tie receptor function and their use (ii)
JOURNAL Patent: EP 1166799-A 24 02-JAN-2002;
SCHERING AKTIEGENSCHAFT (DE)
FEATURES
source 1. .1626
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 68.2%; Score 1599.4; DB 6; Length 1626;
Best Local Similarity 99.9%; Pred. No. 1.6e-276;
Matches 1611; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 734 GGCAATTCCTAGATCTATAGTAGTATCAGGATATATTTTCTTTTAAATATATTTTGG 793
Db 1 GGCAATTCCTAGATCTATAGTAGTATCAGGATATATTTTCTTTTAAATATATTTTGG 60
QY 794 TTATTTTGAATACAGACATTTGCTCAAAATTTTTCATCTTTGCACAATAGTATGACTTTTC 853
Db 61 TTATTTTGAATACAGACATTTGCTCAAAATTTTTCATCTTTGCACAATAGTATGACTTTTC 120
QY 854 ACTAGAACTTCTCAACATTTGGAACTTTGCAAAATATGAGCATCATATGTTTAAAGGCTG 913
Db 121 ACTAGAACTTCTCAACATTTGGAACTTTGCAAAATATGAGCATCATATGTTTAAAGGCTG 180

Qy	914	TATCATTTTAAATGCTATGAGATACATGTGTTTCTCCCTATGCGCAAAACAGGTGAACAAACGT	973
Db	181	TATCATTTTAAATGCTATGAGATACATGTGTTTCTCCCTATGCGCAAAACAGGTGAACAAACGT	240
Qy	974	AGTGTGTTTTTACTGTACTAAATGTTGGCTACCTGTAATTTTATAGTATGCACTGTCA	1033
Db	241	AGTGTGTTTTTACTGTACTAAATGTTGGCTACCTGTAATTTTATAGTATGCACTGTCA	300
Qy	1034	GAATAAGGCAAGACAAATGGCTCTTGTACTGAATATCTTCGGCAAACTTATGGGTCTTC	1093
Db	301	GAATAAGGCAAGACAAATGGCTCTTGTACTGAATATCTTCGGCAAACTTATGGGTCTTC	360
Qy	1094	ATTTTCCTGACAGACAGGATTTGACTCAATATTTGTAGAGCTTGCGTAGAATGGATACAT	1153
Db	361	ATTTTCCTGACAGACAGGATTTGACTCAATATTTGTAGAGCTTGCGTAGAATGGATACAT	420
Qy	1154	GGTGTGTGACACTGGTAGAAATGGTTTTAGTTTTATTTGACTCAGAAATTCATCTCAGGATG	1213
Db	421	GGTGTGTGACACTGGTAGAAATGGTTTTAGTTTTATTTGACTCAGAAATTCATCTCAGGATG	480
Qy	1214	AATCTTTTATGCTTTTTTATTTGTAAGCATATCTGAAATTTACTTTTATAAAGATGGTTTTAG	1273
Db	481	AATCTTTTATGCTTTTTTATTTGTAAGCATATCTGAAATTTACTTTTATAAAGATGGTTTTAG	540
Qy	1274	AAAGCTTTGCTAAATAATTTGGCCTAGGATGGTAACTTCATTTTTCAGTTTGCCAAGGGGT	1333
Db	541	AAAGCTTTGCTAAATAATTTGGCCTAGGATGGTAACTTCATTTTTCAGTTTGCCAAGGGGT	600
Qy	1334	AGAAAAATAATATGTGTGTTGTTTATGTTTAAACATATTTATGAGTACTATCTATGA	1393
Db	601	AGAAAAATAATATGTGTGTTGTTTATGTTTAAACATATTTATGAGTACTATCTATGA	660
Qy	1394	ATGTATTTTAAATATTTTTCATATTTCTGTGCAAGCATTTTAAATTTCCACACAGTGGAGT	1453
Db	661	ATGTATTTTAAATATTTTTCATATTTCTGTGCAAGCATTTTAAATTTCCACACAGTGGAGT	720
Qy	1454	CCATTTAGCCAGTGGGAAAGTCTTGGAACTCAGGTTACCGTTGAAAGATATGCTGCAG	1513
Db	721	CCATTTAGCCAGTGGGAAAGTCTTGGAACTCAGGTTACCGTTGAAAGATATGCTGCAG	780
Qy	1514	CCATCTCTTTGATCTGTGCTTTAAACTGTAAATTTATAGACAGCTAAATCCCTAACTTGGGA	1573
Db	781	CCATCTCTTTGATCTGTGCTTTAAACTGTAAATTTATAGACAGCTAAATCCCTAACTTGGGA	840
Qy	1574	TCTGGAATGCAATTAGTTATGACCTTGTAACCATCCCAAGATTTTCAGGGGCATCGTGGGTT	1633
Db	841	TCTGGAATGCAATTAGTTATG-CCTTGTACCATCCCAAGATTTTCAGGGGCATCGTGGGTT	899
Qy	1634	TGGTCTAGTGTATGAAACCAAGAACAGAGAGATCCAGCTGAAAAAGAGTGATCCTCAA	1693
Db	900	TGGTCTAGTGTATGAAACCAAGAACAGAGAGATCCAGCTGAAAAAGAGTGATCCTCAA	959
Qy	1694	TATCCTAACTAACTGGTCCTCAACTCAAGCAGAGGTTTCTCACTCTGGCACTGTGATCAT	1753
Db	960	TATCCTAACTAACTGGTCCTCAACTCAAGCAGAGGTTTCTCACTCTGGCACTGTGATCAT	1019
Qy	1754	GAACCTTAGTAGGGGATCTGTGTATTTTATACAAATTTAAATACAAATGCTTTACATTTG	1813
Db	1020	GAACCTTAGTAGGGGATCTGTGTATTTTATACAAATTTAAATACAAATGCTTTACATTTG	1079
Qy	1814	ATAAAATTTCTTAAAGACAAATCTGCATTTTATTTCTGCATCCCATTCATCATATTA	1873
Db	1080	ATAAAATTTCTTAAAGACAAATCTGCATTTTATTTCTGCATCCCATTCATCATATTA	1139
Qy	1874	GAACTTAAGATATTTATCTATGAAGATATAAATGGTGCAGAGAGACTTTTCATCTGTGATT	1933
Db	1140	GAACTTAAGATATTTATCTATGAAGATATAAATGGTGCAGAGAGACTTTTCATCTGTGATT	1199
Qy	1934	CGGTTGTTTCTTAGGGTTCTTAGCACTGTATGCCCTGCACAGCATGTGATGTGAAATAA	1993
Db	1200	CGGTTGTTTCTTAGGGTTCTTAGCACTGTATGCCCTGCACAGCATGTGATGTGAAATAA	1259
Qy	1994	AATGGATTTCTCTATPAGTAAATGAGTTTCCCTCTGGGAGAGGTTCTGGTACTGCAATCAC	2053

Db	1260	AATGGGATTCCTTATAGCTAAATGAGTCCCTCTGGGAGAGTCTGGTACTGCAATCAC	1319
Qy	2054	AATGCCAGATGGTGTATATGGGCTATTTCGTGAAGTAAGTGGAATGCTATGAAGTAA	2113
Db	1320	AATGCCAGATGGTGTATATGGGCTATTTCGTGAAGTAAGTGGAATGCTATGAAGTAA	1379
Qy	2114	GTGTGTTGGTTTCATCTTATGAAACTCTTGAATGCAATGGCTTTGTATGGAATAAAT	2173
Db	1380	GTGTGTTGGTTTCATCTTATGAAACTCTTGAATGCAATGGCTTTGTATGGAATAAAT	1439
Qy	2174	TTGTGCAATATGATGTCATCAACTTTGCAATGAAATGAAATTTGGTTGCTATTTATATG	2233
Db	1440	TTGGTGCATATGATGTCATCAACTTTGCAATGAAATTTGGTTGCTATTTATATG	1499
Qy	2234	TATTATACCTGTACGGCTTCTAGTTGCTTCAACATTTTATTAACCATTTTGTACATATT	2293
Db	1500	TATTATACCTGTACGGCTTCTAGTTGCTTCAACATTTTATTAACCATTTTGTACATATT	1559
Qy	2294	TTACTTGAAATATTTTAAATGAAATTTAAATAAACAATTTGATAGTTTACAT	2346
Db	1560	TTACTTGAAATATTTTAAATGAAATTTAAATAAACAATTTGATAGTTTACAT	1612
RESULT 8			
AX441124		1626 bp	linear
LOCUS	AX441124	1626 bp	DNA
DEFINITION	Sequence 24 from Patent WO0197850.		PAT 28-JUN-2002
ACCESSION	AX441124		
VERSION	AX441124.1	GI:21665719	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Siemeister, G., Haberey, M. and Thierauch, K.H.		
TITLE	Combinations and compositions which interfere with vegf/vegfr and angiotensin/tie receptor function and their use (il)		
JOURNAL	Patent: WO 0197850-A 24 27-DEC-2001; SCHERING AKTIEGESELLSCHAFT (DE) ; Siemeister, Gerhard (DE) ; Haberey, Martin (DE) ; Thierauch, Karl-Heinz (DE)		
FEATURES	Location/Qualifiers		
source	1..1626		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	68.2%;	Score 1599.4;	DB 6; Length 1626;
Best Local Similarity	99.9%;	Pred. No. 1.6e-276;	
Matches 1611;	Conservative 0;	Mismatches 1;	Indels 1; Gaps 1;
Qy	734	GGACAATTCCTAGAACTCTATAGTATGATCAGGATATATTTTGGCTTTAAATATATTTGG	793
Db	1	GGACAATTCCTAGAACTCTATAGTATGATCAGGATATATTTTGGCTTTAAATATATTTGG	60
Qy	794	TTATTTTGAATACAGACATTTGGCTCCAAATTTTCATCTTTGCACAAATAGTATGACTTTTC	853
Db	61	TTATTTTGAATACAGACATTTGGCTCCAAATTTTCATCTTTGCACAAATAGTATGACTTTTC	120
Qy	854	ACTAGAACTTCTCAACATTTTGGGAACCTTTGGCAATATGAGCATCATATGTGTTAAGCTG	913
Db	121	ACTAGAACTTCTCAACATTTTGGGAACCTTTGGCAATATGAGCATCATATGTGTTAAGCTG	180
Qy	914	TATCATTTAATGCTATGAGATACATTTGTTTCTCCCTATGCGCAACAGGTGCAACAAACGT	973
Db	181	TATCATTTAATGCTATGAGATACATTTGTTTCTCCCTATGCGCAACAGGTGCAACAAACGT	240
Qy	974	AGTTGTTTTTACTGATACATAAATGTTGGCTACCTGTGATTTTATAGTATGCAATGTCA	1033
Db	241	AGTTGTTTTTACTGATACATAAATGTTGGCTACCTGTGATTTTATAGTATGCAATGTCA	300

QY	1034	GAAAAAGGCAAGACAAATGGCCCTTTGTACTGAATACTTCGGCAAACCTTANTGGGTCTTC	1091
Db	301	GAAAAAGGCAAGACAAATGGCCCTTTGTACTGAATACTTCGGCAAACCTTANTGGGTCTTC	360
QY	1094	ATTTTCTGCAGACAGGATTTGACTCATAATTTTGTAGAGCTTCGGTAGAATGATTACAT	1153
Db	361	ATTTTCTGCAGACAGGATTTGACTCATAATTTTGTAGAGCTTCGGTAGAATGATTACAT	420
QY	1154	GGTAGTGATGCACTGGTAGAAAATGGTTTTAGTTATTGACTCAGAATTCATCTCAGGATG	1213
Db	421	GGTAGTGATGCACTGGTAGAAAATGGTTTTAGTTATTGACTCAGAATTCATCTCAGGATG	480
QY	1214	AATCTTTTATGTCCTTTTATTGTAAGCATATCTGAATTTTACTTTATAAAGATCGTTTTAG	1273
Db	481	AATCTTTTATGTCCTTTTATTGTAAGCATATCTGAATTTTACTTTATAAAGATCGTTTTAG	540
QY	1274	AAAGCTTTGTCTAAAAATTTGGCCTAGGAATGGTAACCTTCATTTCTAGTTGCCAAGGGGT	1333
Db	541	AAAGCTTTGTCTAAAAATTTGGCCTAGGAATGGTAACCTTCATTTCTAGTTGCCAAGGGGT	600
QY	1334	AGAAAAATATATGTGTGTTTATGTTATGTTTAAACATATATTAGGTACTATCTATGA	1393
Db	601	AGAAAAATATATGTGTGTTTATGTTATGTTTAAACATATATTAGGTACTATCTATGA	660
QY	1394	ATGTATTTAAATATTTTTCATATTTCTGTACAAGCAATTTAATTTGCCAACAGTGGAGT	1453
Db	661	ATGTATTTAAATATTTTTCATATTTCTGTACAAGCAATTTAATTTGCCAACAGTGGAGT	720
QY	1454	CCATTTAGCCGAGTGGGAAAGTCTWTGAAACTCAGGTTACCTTGTAAGAGATATGCTGGCAG	1513
Db	721	CCATTTAGCCGAGTGGGAAAGTCTWTGAAACTCAGGTTACCTTGTAAGAGATATGCTGGCAG	780
QY	1514	CCATCTCTTTGATCTGTCTTAAACTGTAAATTTATAGCACGCTPAAATCCCTAACTTGGGA	1573
Db	781	CCATCTCTTTGATCTGTCTTAAACTGTAAATTTATAGCACGCTPAAATCCCTAACTTGGGA	840
QY	1574	TCTGGAATGCAATTAGTTTATGACCTTGTACCATTCCCAGAAATTTCCAGGGGCATCGTGGGTT	1633
Db	841	TCTGGAATGCAATTAGTTTATG-CCTTGTACCAATCCCAGAAATTTCCAGGGGCATCGTGGGTT	899
QY	1634	TGGTCTAGTGNTTGAAACACAAGACAGAGAGATCCAGCTGAAAAAGAGTGATCCTCAA	1693
Db	900	TGGTCTAGTGNTTGAAACACAAGACAGAGAGATCCAGCTGAAAAAGAGTGATCCTCAA	959
QY	1694	TATCCTAACTAACTGGTCCCTCAACTCAAGCAGAGTTCCTCACCTCGGCACCTGTGATCAT	1753
Db	960	TATCCTAACTAACTGGTCCCTCAACTCAAGCAGAGTTCCTCACCTCGGCACCTGTGATCAT	1019
QY	1754	GAACTTTAGTAGAGGGGATTTGTGTGTATTTTATACAAATTTTATACAAATGCTTTACATTG	1813
Db	1020	GAACTTTAGTAGAGGGGATTTGTGTGTATTTTATACAAATTTTATACAAATGCTTTACATTG	1079
QY	1814	ATAAAAATCTTAAAGACMAAATCGCAATTTTATTTCTTGCACTCCACATTCCAATCATATTA	1873
Db	1080	ATAAAAATCTTAAAGACMAAATCGCAATTTTATTTCTTGCACTCCACATTCCAATCATATTA	1139
QY	1874	GAACTAGATATTTATCTATGAAGATATAAATGGTGCAGAGAGACTTTTCATCTGTGGATT	1933
Db	1140	GAACTAGATATTTATCTATGAAGATATAAATGGTGCAGAGAGACTTTTCATCTGTGGATT	1199
QY	1934	CGCTTGTTCTTTAGGGTTCTTAGACATGATCGCTGCCACAAGCATGTGATATGCAAAATAA	1993
Db	1200	CGCTTGTTCTTTAGGGTTCTTAGACATGATCGCTGCCACAAGCATGTGATATGCAAAATAA	1259
QY	1994	AATGGATTCCTTATAGCTAAATGAGTTCCCTCTGGGAGAGTTCTGGTACTGCAATCAC	2053
Db	1260	AATGGATTCCTTATAGCTAAATGAGTTCCCTCTGGGAGAGTTCTGGTACTGCAATCAC	1319
QY	2054	AATGCCAGATGGTCTTTATGGGCTATTTGTGTAGTAGTGGTAGAGATGCTATGAGTAA	2113
Db	1320	AATGCCAGATGGTCTTTATGGGCTATTTGTGTAGTAGTGGTAGAGATGCTATGAGTAA	1379
QY	2114	GTGTGTGTTGTTTTCACTTATGGAACCTCTTGATGCATGTGCTTTTGTATGGAATAAATT	2173

Db	1380	GTGTTGTTGTTTCATCTTTATGGAACACTCTTGAAGCATGTGCTTTTGTATGGAAATAATT	1439
Qy	2174	TTGGTGCAGATATGATGTCATTTCAACTTTGTCATTGGAATTTGATTTGGTTCATTTATATG	2233
Db	1440	TTGGTGCAGATATGATGTCATTTCAACTTTGTCATTGGAATTTGATTTGGTTCATTTATATG	1499
Qy	2234	TATATACCTCTCAACGCTTCTAGTTGCTTCAACCACTTTATACCACTTTTGTGTACATATT	2293
Db	1500	TATATACCTCTCAACGCTTCTAGTTGCTTCAACCACTTTATACCACTTTTGTGTACATATT	1559
Qy	2294	TTACTTGAAAATATTTTAAATGGAATTTTAAATAAACATTTGATAGTTTACAT	2346
Db	1560	TTACTTGAAAATATTTTAAATGGAATTTTAAATAAACATTTGATAGTTTACAT	1612
<p> RESULT 9 AX351446/c LOCUS DEFINITION Sequence 193 from Patent WO0196390. ACCESSION AX351446 VERSION AX351446.1 GI:18616793 KEYWORDS SOURCE Homo sapiens (human) ORGANISM Homo sapiens REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. TITLE 1 JOURNAL Jiang, Y.; Hepler, W.T.; Clapper, J.D.; Wang, A. and Secrist, H. FEATURES Compositions and methods for the therapy and diagnosis of colon source cancer ORIGIN Patent: WO 0196390-A 193 20-DEC-2001; CORIXA CORPORATION (US) Location/Qualifiers 1..586 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" </p>			
<p> Query Match 24.8%; Score 582.8; DB 6; Length 586; Best Local Similarity 99.7%; Pred. No. 2e-94; Matches 584; Conservative 0; Mismatches 2; Indels 0; Gaps 0; </p>			
Qy	476	AACCTTCTTATGCTCTGAAAGAAAGAAAGCAATGTAATCACTATGATTTATGCAAAA	535
Db	586	AACCTTCTTATGCTCTGAAAGAAAGAAAGCAATGTAATCACTATGATTTATGCGAAA	527
Qy	536	CAACCAAGATTTCTCCACAAATTTTAACTATATCTGATCCTCTTCTTGAGAGAAATTTGTAC	595
Db	526	CAACCAAGATTTCTCCACAAATTTTAACTATATCTGATCCTCTTCTTGAGAGAAATTTGTAC	467
Qy	596	CTAATAGTTTTTCTTATGAATGTTATTTACTTCTGTTATATAAATCTATATAATTT	655
Db	466	CTAATAGTTTTTCTTATGAATGTTATTTACTTCTGTTATATAAATCTATATAATTT	407
Qy	656	CCTACTTAAGTCCTTAAGAACTGGGTTCTTCTCTTGATGTTATTCATCTTCAGAAAGAAA	715
Db	406	CCTACTTAAGTCCTTAAGAACTGGGTTCTTCTCTTGATGTTATTCATCTTCAGAAAGAAA	347
Qy	716	CAACACTTTTACTCTTTTAGGACAAATTCCTAGAATCTATAGTAGTATCAGGATATATTTTG	775
Db	346	CAACACTTTTACTCTTTTAGGACAAATTCCTAGAATCTATAGTAGTATCAGGATATATTTTG	287
Qy	776	CTTTAAATATATTTTGGTTATTTTGAATACAGACATTTGGCTCCAAATTTTCATCTTTTGC	835
Db	286	CTTTAAATATATTTTGGTTATTTTGAATACAGACATTTGGCTCCAAATTTTCATCTTTTGC	227
Qy	836	ACAATAGTATGACATTTTCACTTAGAACCTTCTCAACATTTGGGAATTTGGCAATATGACGA	895
Db	226	ACAATAGTATGACATTTTCACTTAGAACCTTCTCAACATTTGGGAATTTGGCAATATGACGA	167
Qy	896	TCATATGTGTTTAAAGCGTGATCATTTTAAATGCTATGAGATACATTTTCTCCCATGCCC	955

```
Db 166 TCATATGTGTAAAGCTGTATCATTTTAATGCTATGAGATACATGTTGTTTCTCCCTATGCC 107
Qy 956 AAACAGGTGAACAAACGCTAGTCTGTTTCTTACTGATACTAAAGTGTGGCTACCTGTGATTT 1015
Db 106 AAACAGGTGAACAAACGCTAGTCTGTTTCTTACTGATACTAAAGTGTGGCTACCTGTGATTT 47
Qy 1016 TATAGTATGCACATGTCTAGAAAAGGCAAGCAAAATGGGCTCTTGT 1061
Db 46 TATAGTATGCACATGTCTAGAAAAGGCAAGCAAAATGGGCTCTTGT 1

RESULT 10
LOCUS AR273526 675 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 1269 from patent US 6504010.
ACCESSION AR273526
VERSION AR273526.1 GI:29705411
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 675)
Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S.,
Carter,D., Retter,W.M., Mannion,J. and Fan,L.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: US 6504010-A 1269 07-JAN-2003;
FEATURES
source Location/Qualifiers
1..675
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 24.8%; Score 581.4; DB 6; Length 675;
Best Local Similarity 97.3%; Pred. No. 3.4e-94;
Matches 622; Conservative 0; Mismatches 13; Indels 4; Gaps 3;
Qy 1673 CTGAAAAGAGTGATCTCAATATCCTAACTAACTGGTCTCACTCAAGCAGAGTTTCT 1732
Db 1 CTGAAAAGAGTGATCTCAATATCCTAACTAACTGGTCTCACTCAAGCAGAGTTTCT 60
Qy 1733 TCACCTCTGGCAGCTGATCATGAACTTAGTAGAGGGGATGTTGTATTTTATACAAT 1792
Db 61 TCACCTCTGGCAGCTGATCATGAACTTAGTAGAGGGGATGTTGTATTTTATACAAT 120
Qy 1793 TTAATACAATGCTTCAATGATAAAATCTTAAAGAGCAAACTGCATTTTATTTCTGC 1852
Db 121 TTAATACAATGCTTCAATGATAAAATCTTAAAGAGCAAACTGCATTTTATTTCTGC 180
Qy 1853 ATCCACATTCCAATCATATTAGAACTAAGATATTTATCTATGAAGATATAAATGGTGCAG 1912
Db 181 ATCCACATTCCAATCATATTAGAACTAAGATATTTATCTATGAAGATATAAATGGTGCAG 240
Qy 1913 AGAGACTTTCAATCTGTGGATGGGTTGTTTCTTAGGGTTCTTAGGCTAGTGCCTGCACA 1972
Db 241 AGAGACTTTCAATCTGTGGATGGGTTGTTTCTTAGGGTTCTTAGGCTAGTGCCTGCACA 300
Qy 1973 AGCATGTGATGTGAATAAATGGAATCTTCTATAGCTAAATGAGTCCCTCTGGGGA 2032
Db 301 AGCATGTGATGTGAATAAATGGAATCTTCTATAGCTAAATGAGTCCCTCTGGGGA 360
Qy 2033 GAGTTCTGGTACTGCAATCAATGCCAGATGGTGTATGGGCTATTTGTGTAAAGTAAG 2092
Db 361 GAGTTCTGGTACTGCAATCAATGCCAGATGGTGTATGGGCTATTTGTGTAAAGTAAG 420
Qy 2093 TGGTAAAGATGCTATGAAGTAAGTGTGTTTGTGTTTTCATCTTATGAAACTCTTGATGCATG 2152
Db 421 TGGTAAAGATGCTATGAAGTAAGTGTGTTTGTGTTTTCATCTTATGAAACTCTTGATGCATG 480
Qy 2153 TGCCTTTGTATGGAATAAATTTTGGTCAATATGATGTCATTCATCACTTTGCATTTGAATTG 2212
Db 481 TGCCTTTGTATGGAATAAATTTTGGTCAATATGATGTCATTCATCACTTTGCATTTGAATTG 540
```

```
Qy 2213 -AATTTGGTGTATTTATATGATTTATA-CCTGTCAAGCTTCTAGTTGCTTCAACCAATT 2270
Db 541 AAAATTTGGGTGGAATTTATATGATTTATATACCTGTCAAGCTTCTAGTTGCTTCAACCAATT 600
Qy 2271 TTATAAACCAATTTTGTACATATTTTACTTGAAATATTTT 2309
Db 601 T--ATACCATTTTGNACATATTTTACTTGNAAATATTT 637

RESULT 11
LOCUS AR277107 675 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 1269 from patent US 6509448.
ACCESSION AR277107
VERSION AR277107.1 GI:29710754
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 675)
Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S.,
Carter,D., Retter,W.M., Mannion,J., Fan,L. and Wang,A.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: US 6509448-A 1269 21-JAN-2003;
FEATURES
source Location/Qualifiers
1..675
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 24.8%; Score 581.4; DB 6; Length 675;
Best Local Similarity 97.3%; Pred. No. 3.4e-94;
Matches 622; Conservative 0; Mismatches 13; Indels 4; Gaps 3;
Qy 1673 CTGAAAAGAGTGATCTCAATATCCTAACTAACTGGTCTCACTCAAGCAGAGTTTCT 1732
Db 1 CTGAAAAGAGTGATCTCAATATCCTAACTAACTGGTCTCACTCAAGCAGAGTTTCT 60
Qy 1733 TCACCTCTGGCAGCTGATCATGAACTTAGTAGAGGGGATGTTGTATTTTATACAAT 1792
Db 61 TCACCTCTGGCAGCTGATCATGAACTTAGTAGAGGGGATGTTGTATTTTATACAAT 120
Qy 1793 TTAATACAATGCTTCAATGATAAAATCTTAAAGAGCAAACTGCATTTTATTTCTGC 1852
Db 121 TTAATACAATGCTTCAATGATAAAATCTTAAAGAGCAAACTGCATTTTATTTCTGC 180
Qy 1853 ATCCACATTCCAATCATATTAGAACTAAGATATTTATCTATGAAGATATAAATGGTGCAG 1912
Db 181 ATCCACATTCCAATCATATTAGAACTAAGATATTTATCTATGAAGATATAAATGGTGCAG 240
Qy 1913 AGAGACTTTCAATCTGTGGATGGGTTGTTTCTTAGGGTTCTTAGGCTAGTGCCTGCACA 1972
Db 241 AGAGACTTTCAATCTGTGGATGGGTTGTTTCTTAGGGTTCTTAGGCTAGTGCCTGCACA 300
Qy 1973 AGCATGTGATGTGAATAAATGGAATCTTCTATAGCTAAATGAGTCCCTCTGGGGA 2032
Db 301 AGCATGTGATGTGAATAAATGGAATCTTCTATAGCTAAATGAGTCCCTCTGGGGA 360
Qy 2033 GAGTTCTGGTACTGCAATCAATGCCAGATGGTGTATGGGCTATTTGTGTAAAGTAAG 2092
Db 361 GAGTTCTGGTACTGCAATCAATGCCAGATGGTGTATGGGCTATTTGTGTAAAGTAAG 420
Qy 2093 TGGTAAAGATGCTATGAAGTAAGTGTGTTTGTGTTTTCATCTTATGAAACTCTTGATGCATG 2152
Db 421 TGGTAAAGATGCTATGAAGTAAGTGTGTTTGTGTTTTCATCTTATGAAACTCTTGATGCATG 480
Qy 2153 TGCCTTTGTATGGAATAAATTTTGGTCAATATGATGTCATTCATCACTTTGCATTTGAATTG 2212
Db 481 TGCCTTTGTATGGAATAAATTTTGGTCAATATGATGTCATTCATCACTTTGCATTTGAATTG 540
Qy 2213 -AATTTGGTGTATTTATATGATTTATA-CCTGTCAAGCTTCTAGTTGCTTCAACCAATT 2270
```

||||| 541 AAATTTGGGTGATTTATATGTTATATACCGTGCACGCTTAGTGTCTTCAACCAAT 600
QY 2271 TTATAACCAATTTTGTACATATTTTACCTGAAATATTT 2309
Db 601 T--ATACCAATTTTGNACATATTTTACTTGNAAATATTT 637

RESULT 12
AX368559
LOCUS AX368559 675 bp DNA linear PAT 16-FEB-2002
DEFINITION Sequence 1269 from Patent WO0204514.
ACCESSION AX368559
VERSION AX368559.1 GI:18856632
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Wang, T., Watanabe, Y., Henderson, R.A., Johnson, J.C., Retter, M.W.,
AUTHORS Marnerakis, M., Carter, D., Fanger, G.R., Vedvick, T.S., Bangur, C.S.,
McNabb, A., Fanger, N., Switzer, A., McNeill, P.D. and Clapper, J.D.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0204514-A 1269 17-JAN-2002;
CORIXA CORPORATION (US)
FEATURES
source
1. 675
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 24.8%; Score 581.4; DB 6; Length 675;
Best Local Similarity 97.3%; Pred. No. 3.4e-94;
Matches 622; Conservative 0; Mismatches 13; Indels 4; Gaps 3;

QY 1673 CTGAAAGAGTGATCCTCAATATCTTAATCTTAAGTGGTCTCAAGCAGAGTTCT 1732
Db 1 CTGAAAGAGTGATCCTCAATATCTTAATCTTAAGTGGTCTCAAGCAGAGTTCT 60

QY 1733 TCACCTGCGCATGATCATGAATCTTAGTAGAGGGGATGTGTGATATTTATACAAT 1792
Db 61 TCACCTGCGCATGATCATGAATCTTAGTAGAGGGGATGTGTGATATTTATACAAT 120

QY 1793 TTAATACATGCTTACATGTGATTAATCTTAAGAGCAAACTGATTTTATTCGC 1852
Db 121 TTAATACATGCTTACATGTGATTAATCTTAAGAGCAAACTGATTTTATTCGC 180

QY 1853 ATCCATTCCTCAATCATATTAAGACTAAGATATTTATCTAAGATATAAATGGTGAG 1912
Db 181 ATCCATTCCTCAATCATATTAAGACTAAGATATTTATCTAAGATATAAATGGTGAG 240

QY 1913 AGAGATTTTCATCTGTTGGATCGGTGTTCTTAGGTTCTTAGCAGCTGATGCTGCACA 1972
Db 241 AGAGATTTTCATCTGTTGGATCGGTGTTCTTAGGTTCTTAGCAGCTGATGCTGCACA 300

QY 1973 AGCATGTGATATGTGAAATATAAATGATTCCTTATAGCTAATAGATTCCTCTGGGGA 2032
Db 301 AGCATGTGATATGTGAAATATAAATGATTCCTTATAGCTAATAGATTCCTCTGGGGA 360

QY 2033 GAGTTCTGATGCTGATCATGATCCAGATCGGTGTTATAGGCTATTTGTGTAGTAAG 2092
Db 361 GAGTTCTGATGCTGATCATGATCCAGATCGGTGTTATAGGCTATTTGTGTAGTAAG 420

QY 2093 TGGTAAGATGCTATGAAGTAAGTGTGTTTGTGTTTTCATCTTATGAAACTCTTTGATGCATG 2152
Db 421 TGGTAAGATGCTATGAAGTAAGTGTGTTTGTGTTTTCATCTTATGAAACTCTTTGATGCATG 480

QY 2153 TGGTTTGTATGGAATAAATTTTGGTGCAATGATGATCAATTCATCTTGCATTTGCAATTTG 2212
Db 481 TGGTTTGTATGGAATAAATTTTGGTGCAATGATGATCAATTCATCTTGCATTTGCAATTTG 540

QY 2213 -AATTTGGTGTGATTTATATGTTATATATA-CCTGTCAACGCTTCTAGTGTCTTCAACCAAT 2270
Db 541 AAATTTGGGTGATTTATATGTTATATATACCGTGCACGCTTAGTGTCTTCAACCAAT 600

QY 2271 TTATAACCAATTTTGTACATATTTTACCTGAAATATTT 2309
Db 601 T--ATACCAATTTTGNACATATTTTACTTGNAAATATTT 637

RESULT 13
AX341176
LOCUS AX341176 586 bp DNA linear PAT 10-JAN-2002
DEFINITION Sequence 1423 from Patent WO0196388.
ACCESSION AX341176
VERSION AX341176.1 GI:18137158
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Jiang, Y., Harlocker, S.L. and Secrist, H.
AUTHORS Compositions and methods for the therapy and diagnosis of colon
TITLE cancer
JOURNAL Patent: WO 0196388-A 1423 20-DEC-2001;
CORIXA CORPORATION (US)
FEATURES
source
1. 586
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 24.2%; Score 568; DB 6; Length 586;
Best Local Similarity 99.0%; Pred. No. 8.7e-92;
Matches 580; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 464 ACTACCCCAATTAACCTTCTTATTGCTCTGAAGAGAGAAAGCAATGTAATCACTATGA 523
Db 1 ACTACCCCAATTAACCTTCTTATTGCTCTGAAGAGAGAAAGCAATGTAATCACTATGA 60

QY 524 TTATTGCAAAACCAACAGAAATCTCAACAATTTTAAAGTAATCTGATCTCTTCTTGA 583
Db 61 TTATTGCGCAACCAACAGAAATCTCAACAATTTTAAAGTAATCTGATCTCTTCTTGA 120

QY 584 GAAATTTGTACCTAATAGTTTTTCTTATGAATGTTATTACTCTGGTATAAATCAAT 643
Db 121 GAAATTTGTACCTAATAGTTTTTCTTATGAATGTTATTACTCTGGTATAAATCAAT 180

QY 644 TTCTATAAATTTCTTAACTTAAGCTTTAAGAACTGGGTTCTTCTTTGATGTTATTCATGT 703
Db 181 TTCTATAAATTTCTTAACTTAAGCTTTAAGAACTGGGTTCTTCTTTGATGTTATTCATGT 240

QY 704 TCAGAAAGGAAACCAACACTTTTACTCTTTTAGGACAATTCCTAGAACTCTATAGTAGTATCA 763
Db 241 TCAGAAAGGAAACCAACACTTTTACTCTTTTAGGACAATTCCTAGAACTCTATAGTAGTATCA 300

QY 764 GATATATTTTGTGTTTAAATATATTTTGGTTATTTTGAATACAGACATTTGGCTCCAAAT 823
Db 301 GATATATTTTGTGTTTAAATATATTTTGGTTATTTTGAATACAGACATTTGGCCCCCAAT 360

QY 824 TTTTCATCTTTGCAACATAGTACCTTTTCACTAGAACTTCTCAACATTTGGGAACCTTG 883
Db 361 TTTTCATCTTTGCAACATAGTACCTTTTCACTAGAACTTCTCAACATTTGGGAACCTTG 420

QY 884 CAAATATGAGCATCATATGTTTAAAGCTG-TATCAATTTAAATGCTATGAGATACATGTT 942
Db 421 CAAATATGAGCATCATATGTTTAAAGCTGTTATCAATTTAAATGCTATGAGATACATGTT 480

QY 943 TTCTCCCTATGCCAAACAGGTGAACAGTCTGTTTCTTTTCTGATCACTAATGTTGG 1002
Db 481 TTCTCCCTATGCCAAACAGGTGAACAGTCTGTTTCTTTTCTGATCACTAATGTTGG 540

QY	1003	CTACCTCTGATTTTATAGTAGTCACATGTCAGAAAAGGCGAGACA	1048
Db	541	CTACCTGNGATTTTATAGTAGTCACATGTCAGAAAAGGCGAGDNA	586
RESULT 14			
AR407382			
LOCUS	AR407382	665 bp	DNA
DEFINITION	Sequence 1269 from patent US 6630574.		linear
ACCESSION	AR407382		
VERSION	AR407382.1	GI:40157193	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 665)		
AUTHORS	Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S., Carter,D., Ketter,M.W., Mannion,J. and Fan,L.		
TITLE	Compositions and methods for the therapy and diagnosis of lung cancer		
JOURNAL	Patent: US 6630574-A 1269 07-OCT-2003;		
FEATURES	Location/Qualifiers		
source	1..665		
ORIGIN	/organism="unknown"		
	/mol_type="genomic DNA"		
Query Match	23.5%;	Score 551.4;	DB 6;
Best Local Similarity	95.8%;	Pred. No. 8e-89;	
Matches	612;	Conservative 0;	Mismatches 13;
			Indels 14;
			Gaps 3;
QY	1673	CTGAAAAGAGTGTACTCTCAATATCCCTAACTAATCGTCTCAACTCAAGCAGAGTTTCT	1732
Db	1	CTGAAAAGAGTGTACTCTCAATATCCCTAACTAATCGTCTCAACTCAAGCAGAGTTTCT	60
QY	1733	TCACTCTGGCACTCTGATCATGAACTTAGTAGAGGGGATGTGTGATTTTATACAAAT	1792
Db	61	TCACTCTGGCACTGTGTATCATGAACCTTAGTAGAGGGGATGTGTGATTTTATACA	117
QY	1793	TTAATACAATGTCCTTACATGATGATAAATCTTAAGAGCAAACTGCATTTTATCTGC	1852
Db	118	-----AATGCTCTTACATGATGATAAATCTTAAGAGCAAACTGCATTTTATCTGC	170
QY	1853	ATCCACATTCCAATCATATTAGAACTAAGATATTTATCTATGAAGATATAAATGGTCAG	1912
Db	171	ATCCACATTCCAATCATATTAGAACTAAGATATTTATCTATGAAGATATAAATGGTCAG	230
QY	1913	AGAGACTTTCACTGTGGATGTCGTGTTCTTATAGGGTTCCTAGCACTGATGCTGCACA	1972
Db	231	AGAGACTTTCACTGTGGATGTCGTGTTCTTATAGGGTTCCTAGCACTGATGCTGCACA	290
QY	1973	AGCATGTGATATGGAATAAATGGAATCTTCTATAGCTAAATAGTTCCTCTCGGGA	2032
Db	291	AGCATGTGATATGGAATAAATGGAATCTTCTATAGCTAAATAGTTCCTCTCGGGA	350
QY	2033	GAGTTCTGGTACTGCAATCAATGCGAGATGGTGTTTATGGGCTATTGTGTAAGTAAG	2092
Db	351	GAGTTCTGGTACTGCAATCAATGCGAGATGGTGTTTATGGGCTATTGTGTAAGTAAG	410
QY	2093	TGTTAAGATGCTATGAAGTAAGTCGTGTTGTTTTTCATCTTATGAAACTCTTGATGCATG	2152
Db	411	TGTTAAGATGCTATGAAGTAAGTCGTGTTGTTTTTCATCTTATGAAACTCTTGATGCATG	470
QY	2153	TGCTTTTGTATGGAATAAATTTTGGTGCAATATGATGTCATTCCACTTTGCAATTAAG	2212
Db	471	TGCTTTTGTATGGAATAAATTTTGGTGCAATATGATGTCATTCCACTTTGCAATTAAG	530
QY	2213	-AATTTTGGTGTATTTATATGTAATTATA-CCTGTCACGGCTTCAGTGTGCTTCAACCAT	2270
Db	531	AAATTTTGGTGTATTTATATGTAATTATA-CCTGTCACGGCTTCAGTGTGCTTCAACCAT	590
QY	2271	TTATAACCAATTTTGTGCATATTTTACTTGGAAATATTTT	2309

```

Db 1694 ATGCTGTTTATGGGCTATTTGTGTAAGTGTGAAGCTATGAAGTAACTGTGTTT 1753
Qy 2122 GTTTTCATCTTATGGAACCTTCATGCACTGCTTTTGTATGGAATAAATTTTGGTGCA 2181
Db 1754 GTTTTCATCTTATGGAACCTTCATGCACTGCTTTTGTATGGAATAAATTTTGGTGCA 1813
Qy 2182 ATATGATGTCATTCAACTTTGGCATTGAATTTTGGTGTATTTATATGATTATAC 2241
Db 1814 ATATGATGTCATTCAACTTTGGCATTGAATTTTGGTGTATTTATATGATTATAC 1873
Qy 2242 CTGTCACGCTTCTAGTTGCTTCAACCACTTTTATTAACCACTTTTGTACATATTTTACTTGA 2301
Db 1874 CTGTCACGCTTCTAGTTGCTTCAACCACTTTTATTAACCACTTTTGTACATATTTTACTTGA 1933
Qy 2302 AAATATTTTAAATGGAATTTTAAATTAACCACTTTGATAGTTTACAT 2346
Db 1934 AAATATTTTAAATGGAATTTTAAATTAACCACTTTGATAGTTTACAT 1978

```

Search completed: April 25, 2004, 18:58:41
 Job time : 6182.51 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 25, 2004, 07:00:48 ; Search time 4507.01 Seconds
(without alignments)

16458.286 Million cell updates/sec

Title: US-10-051-835-18

Perfect score: 2484

Sequence: 1 gttattctgtattataac.....ataaatttttataaaactc 2484

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em_estba.*
2: em_estba.*
3: em_estba.*
4: em_estba.*
5: em_estov.*
6: em_estov.*
7: em_estov.*
8: em_estov.*
9: em_estov.*
10: gb_est2.*
11: gb_est2.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_estom.*
18: em_estom.*
19: em_estom.*
20: em_estom.*
21: em_estom.*
22: em_estom.*
23: em_estom.*
24: em_estom.*
25: em_estom.*
26: em_estom.*
27: em_estom.*
28: gb_est1.*
29: gb_est2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	625.4	25.2	627	13 BX470854	BX470854 DKFZp686B
2	617.6	24.9	778	12 BG676811	BG676811 602623406
3	523.8	21.1	607	9 A180763	A180763 2P43f01.r
4	513	20.7	513	9 AA480136	AA480136 2V42f09.r

C	5	510	20.5	510	10	BE049439	BE049439 xw86ell.x
C	6	504	20.3	639	14	N90350	N90350 Y288e07.r1
C	7	472.8	19.0	504	9	AA031928	AA031928 xk15b08.s
C	8	453	18.2	608	10	AW381442	AW381442 PMO-HT030
C	9	450	18.1	451	9	AI275105	AI275105 ql65d11.x
C	10	442.6	17.8	1201	9	AL541045	AL541045 x152d08.x
C	11	440.8	17.7	452	10	AW130518	AW130518 x152d08.x
C	12	436	17.6	436	9	AI290106	AI290106 qw26d02.x
C	13	434.8	17.5	462	14	N33172	N33172 yv35c11.s1
C	14	433.4	17.4	763	14	CA411843	CA411843 UI-H-E20-
C	15	431	17.4	512	9	AA032029	AA032029 xk15b08.f
C	16	430.4	17.3	761	14	CA426666	CA426666 UI-H-PF1-
C	17	430	17.3	430	9	AA476356	AA476356 wv98a02.f
C	18	427.8	17.2	512	9	AV751313	AV751313 AV751313
C	19	427	17.2	1089	9	AL577084	AL577084 AL577084
C	20	424.8	17.1	764	13	BO772889	BO772889 UI-H-PF0-
C	21	424.4	17.1	765	14	CA423905	CA423905 UI-H-PF1-
C	22	413.4	16.6	415	10	AW613660	AW613660 hh76d05.x
C	23	403.8	16.3	978	13	EX369387	EX369387 EX369387
C	24	389.4	15.7	489	10	BF375154	BF375154 RC3-ST019
C	25	383.6	15.4	507	14	N90723	N90723 za90b12.s1
C	26	382.8	15.4	615	10	BE826206	BE826206 QV4-EN003
C	27	368	14.8	482	9	AA404636	AA404636 zt41e05.f
C	28	362.4	14.6	487	12	BM987655	BM987655 UI-H-CO0-
C	29	350.4	14.1	691	10	AW118825	AW118825 xq97b06.x
C	30	345	13.9	351	9	AM028857	AM028857 wv91e08.x
C	31	341.6	13.8	714	9	AI924840	AI924840 wn23f02.x
C	32	328.4	13.2	658	14	CA415309	CA415309 UI-H-E20-
C	33	324.4	13.1	644	10	AW593610	AW593610 xm03d06.x
C	34	324.2	13.1	657	10	AW993106	AW993106 RC2-BN003
C	35	317.2	12.8	374	9	AI061108	AI061108 an26c04.x
C	36	313	12.6	424	9	AA233637	AA233637 zt43e10.x
C	37	311	12.5	386	10	AW204634	AW204634 UI-H-B11-
C	38	311	12.5	446	14	CB143712	CB143712 K-EST0197
C	39	311	12.5	483	9	AA775662	AA775662 zt26e03.s
C	40	311	12.5	485	9	AI160736	AI160736 qb49c02.x
C	41	311	12.5	490	9	AI204255	AI204255 qv75d04.x
C	42	311	12.5	496	9	AI351758	AI351758 qt03c09.x
C	43	311	12.5	496	10	AW192267	AW192267 xl82c12.x
C	44	311	12.5	543	9	AA975121	AA975121 on06e01.s
C	45	311	12.5	565	9	AI565967	AI565967 tq68g06.x

ALIGNMENTS

627 bp mRNA linear EST 04-SEP-2003
DKFZp686B06123 r1 686 (synonym: hicc3) Homo sapiens cDNA clone
DKFZp686B06123 5', mRNA sequence.

ACCESSION BX470854.1 GI:31665178
VERSION BX470854
KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 627)

BAHR, A., LAUBER, J., MEWES, H.W., WEIL, B., AMID, C., OSANGER, A.,

ROBO, G., HAN, W. and WIEMANN, S. 2003, Mammalian Genome Project, Berlin.

EST (BAHR, A., LAUBER, J., MEWES, H.W., WEIL, B., et al.)

Unpublished (2003)

CONTACT: MIPS

COMMENT

MIPS

Ingolstaedter Landstr. 1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;

sequenced by G. Robo (Hilden/Germany) within the cDNA sequencing

consortium of the German Genome Project.

No sl sequence available.

This clone (DKFZp686B06123) is available at the RZPD in Berlin.

Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: rcapbs@mail.nih.gov
 Tissue Procurement: James Cleaver, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAMI0600 row: k column: 01
 High quality sequence stop: 777.

```

1. 778
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4748664"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP SKn4"
/note="Organ: Skin; Vector: pCMV-Sport6; Site_1: Not I;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5Kb. Library constructed by Life
Technologies. Note: this is a NCI CGAP Library."

```

1	CCTAGAACTGGATTGGCATTTGCTCTTGATCAGATGAGACCTTTGATTATTGGCC	161
4		
2	CCTAGAACTGGATTGGCATTTGCTCTTGATCAGATGAGACCTTTGATTATTGGCC	488
4	CTTCCTTAGGACCTTACACT - CTGTCTTTCTTTGACTTGTCTTTTGTTCCTTCCTTC	171
2	CTTCCTTAGGACCTTACACTACCTGCTTTTCTTTGACTTGTCTTTTGTTCCTTCCTTC	543

```

QY 1743 ATCTAGTCCCTCTTCATGCGATATGTCATTCGCTA-GGTAGAGGTATGCTCTTTATGT 1801
DB 542 ATCTAGTCCCTCTTCATGCGATATGTCATTCGCTA-GGTAGAGGTATGCTCTTTATGT 601
QY 1802 AATGCCACCCGCAATTAGTATTACATAAA-CTTCTCTTTTAAACAATCTGTGCAAGTACAT 1860
DB 602 AATGCCACCCGCAATTAGTATTACATAAAAGCTTCTTTTAAACAATCTGTGCAAGTACAT 661
QY 1861 GCTGCTCTGTTCCATTAGAGATTGACAGAGGTTTCAGTTTGTAGTACTCAAACTTTAT 1920
DB 662 GCTGCTCTGTTCCATTAGAGATTGACAGAGG-TTCAGTTTGTAGTACTCAAACTTTA- 718
QY 1921 TTTAGTCTTGGGAAATCAATTCAGAAATACATCTCTCCAAATCTCTCTTACTCAAA 1980
DB 719 TTTAGTCTTGGGAAATCAATTCAGAAATACATCTCTCCAAATCTCTCTTACTCAAA 778

RESULT 3
AA180763 607 bp mRNA linear EST 10-MAR-1998
LOCUS zp43f01.r1 Stragene muscle 937209 Homo sapiens cDNA clone
DEFINITION IMAGE:612217.5, similar to SW:P43_PANTR P27473 NON-A NON-B
HEPATITIS-ASSOCIATED MICROTUBULAR AGGREGATES PROTEIN ; mRNA
sequence.
ACCESSION AA180763
VERSION AA180763.1 GI:1764237
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 607)
REFERENCE 1
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)
TITLE Contact: Wilson RK
JOURNAL Washington University School of Medicine
COMMENT 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1453 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 386.

FEATURES
source
1..607
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:4644056"
/db_xref="taxon:9606"
/clone="IMAGE:612217"
/tissue_type="muscle"
/dev_stage="adult"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stragene muscle 937209"
/note="Organ: skeletal muscle; Vector: pBluescript SK-;
Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally.
Primer: Oligo dt. Skeletal muscle from patient with
malignant hyperthermia. Average insert size: 1.0 kb;
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG
3' -3' adaptor sequence: 5' CTCGAGTATTTTCTTTTCTTCTTCTGATTCGGGTAGCAATTATTC 1182

QY 1123 GGAAGTCCAGAAACTGGATTGCTCTTCTGACATCTCGGTGGTTCAGTAATTATTC 1182

```

```

DB 1 GGAAGTCCAGAAACTGGATTGCTCTTCTGACATCTCGGTGGTTCAGTAATTATTC 60
QY 1183 CTCTGAGTGGGAGCTGGACCTCTGTAAGAGATGTTCTTAATCTTTCTGCTCGAGCGAAT 1242
DB 61 CTCTGAGTGGGAGCTGGACCTCTGTAAGAGATGTTCTTAATCTTTCTGCTCGAGCGAAT 120
QY 1243 GCTATGGGCTGAGATGACTTCTTAGAGGATTTGCTTTGAGCAAAATAGGTAGATGGTT 1302
DB 121 GCTATGGGCTGAGATGACTTCTTAGAGGATTTGCTTTGAGCAAAATAGGTAGATGGTT 180
QY 1303 TGCTGTGTGGAAGCTTGGAAAGCGTTCAGGTAGTTGGTACTTTCTGCTTGGATCTATTA 1362
DB 181 TGCTGTGTGGAAGCTTGGAAAGCGTTCAGGTAGTTGGTACTTTCTGCTTGGATCTATTA 240
QY 1363 AATACCTGCGACCTCTGCTCTTTTGGTGGTGTGGCTGTGATAGTCTGCTTTT 1422
DB 241 AATACCTGCGACCTCTGCTCTTTTGGTGGTGTGGCTGTGATAGTCTGCTTTT 300
QY 1423 AACCCACTCCCTGGATGCAATTTTCCCTCTTGCATTTCCCTCTTTTCTCGAGTTCATA 1482
DB 301 AACCCACTCCCTGGATGCAATTTTCCCTCTTGCATTTCCCTCTTTTCTCGAGTTCATA 360
QY 1483 CTAGAGAATCTGCATATGTTTTTCCCTTTTGTCTTGAGATGAAGTTTAAATAATC 1542
DB 361 CTAGAGAATCTGCATATGTTTTTCCCTTTTGTCTTGAGATGAAGTTTAAATAATC 420
QY 1543 CACCTCTGTCAATTCACACTCTCTGAACATCCCAAGCTGTATCCCTGGCTCTTTTCTCAG 1602
DB 421 CACCTCTGTCAATTCACACTCTCTGAACATCCCAAGCTGTAT-CTTGGCTCTTTTCTCAG 479
QY 1603 ACTATGTTTCTTACTTGGGACCTAGAACTGGATGGATGGCAATGCTCTCTGATCAGT 1662
DB 480 ACTAAGTTTCTTACTTGGGACCTAGAACTGGATGGATGGCA-TGCTCCNGATCAGAT 538
QY 1663 GAGACCTTTGATTTATTTGGCTTCTCTTAGGACCTTACACTCTCTTCTTTTCTGACTTG 1722
DB 539 GAGACCTT---GAATATTGCCCTCTCTTAAGACCTTA-ANTCCCGCTTCTTTGACTGC 594
QY 1723 CCTTT 1727
DB 595 CTTTT 599

RESULT 4
AA480136 513 bp mRNA linear EST 08-AUG-1997
LOCUS zv42f09.r1 Soares ovary tumor NboHt Homo sapiens cDNA clone
DEFINITION IMAGE:756329.5, mRNA sequence.
ACCESSION AA480136
VERSION AA480136.1 GI:2208287
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 513)
REFERENCE 1
AUTHORS Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,
Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-Merck EST Project 1997
Unpublished (1997)
TITLE Contact: Wilson RK
JOURNAL Washington University School of Medicine
COMMENT 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28M13 rev2 ET from Amersham
High quality sequence stop: 493.

```


N90350 639 bp mRNA linear EST 03-APR-1996
 LOCUS Y288e07.r1 Soares multiple sclerosis 2NBHMP Homo sapiens cDNA
 DEFINITION clone IMAGE:290148 5', mRNA sequence.
 N90350
 ACCESSION N90350.1 GI:1443677
 VERSION
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 639)
 AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohldmann, P. and Wilson, R.
 TITLE The WashU-Merck EST Project
 JOURNAL Unpublished (1995)
 COMMENT Contact: Wilton RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: Enprimer
 High quality sequence stop: 408.
 FEATURES
 Location/Qualifiers
 1..639
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:3905988"
 /db_xref="taxon:9606"
 /clone="IMAGE:290148"
 /sex="male"
 /tissue_type="multiple sclerosis lesions"
 /dev_stage="Age 46"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares multiple sclerosis 2NBHMP"
 /notes="Vector: pT73D (Pharmacia) with a modified polylinker V.TPE; phagemid; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5']
 TGTACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTTTTTT 3',
 double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. RNA from 4 multiple sclerosis lesions from one patient was kindly provided by Dr. Kevin G. Becker (NINDS/NIH)."
 ORIGIN
 Query Match 20.3%; Score 504; DB 14; Length 639;
 Best Local Similarity 92.9%; Pred.No. 5.6e-87;
 Matches 560; Conservative 0; Mismatches 35; Indels 8; Gaps 3;
 QY 1821 ATTACATAAACTTTCTTTTAAACAACTGTGCATAGTACATGCTGCTCTTCCATTAGA 1880
 DB 1 ATTACATAAACTTTCTTTTAAACAACTGTGCATAGTANATGCTGCTCTTCCATTAGA 60
 QY 1881 GATTGACAGAGTTTCAGTTTGTAGTACTCAATCTTATTTTGTAGTCTTGGAAATCAA 1940
 DB 61 GATTGACAGAGTTTCAGTTTGTAGTACTCAATCTTATTTTGTAGTCTTGGAAATCAA 120
 QY 1941 TTCAGAAATACATCTCTCCAAATCTCTCTTACTCAAAATGCTGGAAATCTCATGT 2000
 DB 121 TTCAGAAATACATCTCTCCAAATCTCTCTTACTCAAAATGCTGGAAATCTCATGT 180
 QY 2001 TACTAACTTTGTGCTCTAACTCTGCCATCTTGCTTCCCATCCCTCTCTCTCTCATG 2060

Db 181 TACTAACTTTGTGCTCTAACTCTGCCATCTTGCTTCCCATCCCTCTCTCTCATG 240
 QY 2061 GTACGTGTGCTCTTAATATTAGCGTTGGTTGAGATTTTCAGTGGTCAATATTCTCTTC 2120
 Db 241 GTACGTGTGCTCTTAATATTAGCGTTGGTTGAGATTTTCAGTGGTCAATATTCTCTTC 300
 QY 2121 CCTCTGTTGCTCTTCTCTGAGATAATCCACTAGATAATTTTGTGTTCTTTTCTCAGGG 2180
 Db 301 CCTCTGTTGCTCTTCTCTGAGATAATCCACTAGATAATTTTGTGTTCTTTCTCTC 357
 QY 2181 AATCTAAGGAGGAAATATTCAACTGTGCACAAAGGAAAAAATAGATATGTGAAGGTTT 2240
 Db 358 GAATTAAGGAGGAAATATTCAACTGTGCACAAAGGAAAAAATAGATATGTGAAGGTTT 417
 QY 2241 ACCTAAATTTCTCATCATCACAGAAGTAAATTCAGAAAGGAAAAACACAGACCAA 2300
 Db 418 ACCTAAATTTCTCATCATCACAGAAGTAAATTCAGAAAGGAAAAACACAGACCAA 477
 QY 2301 GAGAAGTATCTAAGACCAAGGGATGTTTATTATTATGTTCTAGGATGAAGAAATGCATA 2360
 Db 478 GAGAAGTANTAAGACCAAGGGATGTTTATTATTATGTTCTAGGATGAAGAAATGCATA 536
 QY 2361 GAACATTTGTAGTACTGTAAATACTAGATAATCAATGATTAGTCAATATTGTGAATA 2420
 Db 537 GA-----CATTTGTAGTCTTGTAAATACTAGATAATCAATGATTAGGCCATAATGTGAATA 592
 QY 2421 TAA 2423
 Db 593 AAA 595
 RESULT 7
 AA031928 504 bp mRNA linear EST 09-MAY-1997
 LOCUS IMAGE.470583 3', mRNA sequence.
 DEFINITION zk15b08.s1 Soares pregnant uterus_NbHPU Homo sapiens cDNA clone
 ACCESSION AA031928
 VERSION AA031928.1 GI:1501891
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 504)
 AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B., Chisoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevasakis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Marra, M.
 TITLE Generation and analysis of 280,000 human expressed sequence tags
 JOURNAL Genome Res. 6 (9), 807-828 (1996)
 MEDLINE 97044478
 PUBMED 8889549
 COMMENT Contact: Wilton RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 670 Std Error: 0.00
 Seq primer: -40M13 fwd from Amersham
 High quality sequence stop: 446.
 FEATURES
 Location/Qualifiers
 1..504
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:3756585"
 /db_xref="taxon:9606"